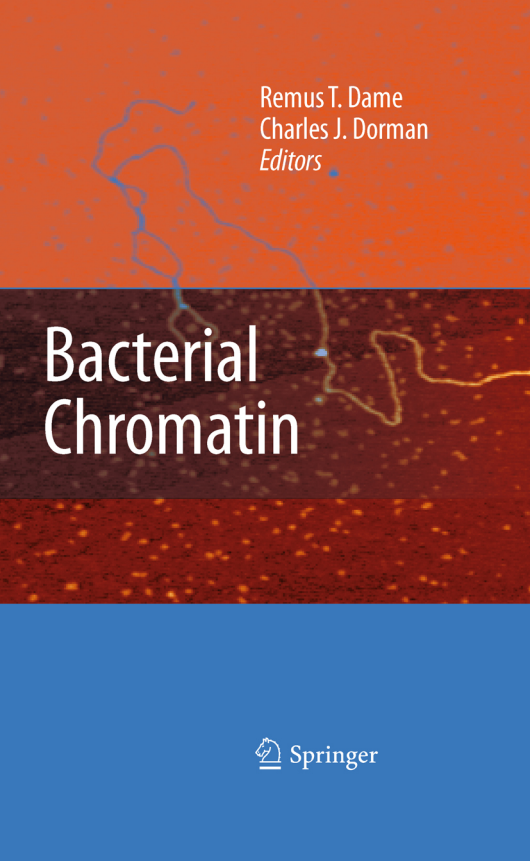




Remus T. Dame
Charles J. Dorman
Editors

Bacterial Chromatin

 Springer

VISIT...

LANZAROTE
Caliente.COM

Bacterial Chromatin

Remus T. Dame • Charles J. Dorman
Editors

Bacterial Chromatin

 Springer

Editors

Remus T. Dame
Faculty of Mathematics and Natural
Sciences, Leiden Institute of Chemistry
Laboratory of Molecular Genetics
Einsteinweg 55
2333 CC, Leiden, Netherlands
and
Faculty of Science
Division of Physics and Astronomy
Section Physics of Complex Systems
VU University Amsterdam
De Boelelaan 1081
1081 HV, Amsterdam, The Netherlands
rtdame@chem.leidenuniv.nl

Charles J. Dorman
Department of Microbiology
School of Genetics and Microbiology
University of Dublin
Trinity College
Dublin 2
Ireland
cjdorman@tcd.ie

ISBN 978-90-481-3472-4 e-ISBN 978-90-481-3473-1

DOI 10.1007/978-90-481-3473-1

Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009941800

© Springer Science+Business Media B.V. 2010

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Cover illustration: H-NS-DNA complex visualized using scanning force microscopy (Courtesy of R.T. Dame)

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The birth and the development of molecular biology and, subsequently, of genetic engineering and biotechnology cannot be separated from the advancements in our knowledge of the genetics, biochemistry and physiology of bacteria and bacteriophages. Also most of the tools employed nowadays by biotechnologists are of bacterial (or bacteriophage) origin and the playground for most of the DNA manipulations still remains within bacteria. The relative simplicity of the bacterial cell, the short generation times, the well defined and inexpensive culturing conditions which characterize bacteria and the auto-catalytic process whereby a wealth of in-depth information has been accumulated throughout the years have significantly contributed to generate a large number of knowledge-based, reliable and exploitable biological systems.

The subtle relationships between phages and their hosts have produced a large amount of information and allowed the identification and characterization of a number of components which play essential roles in fundamental biological processes such as DNA duplication, recombination, transcription and translation. For instance, to remain within the topic of this book, two important players in the organization of the nucleoid, FIS and IHF, have been discovered in this way. Indeed, it is difficult to find a single fundamental biological process whose structural and functional aspects are better known than in bacteria.

However, a notable exception is represented by the physical and functional organization of the bacterial genome. Although some bacteria contain more than one chromosome and some chromosomes are known to be linear, the majority of bacterial cells contain a single circular chromosome. The chromosome of *Escherichia coli* consists of about 4.6 million bp corresponding to a fully extended circumference of about 1.6 mm and rapidly growing bacteria may contain up to almost four genomic equivalents. Thus, the need for compaction of this genetic material to fit within an approximately 500-fold smaller volume is obvious; likewise, also clear is the need for a dynamic “chromatin” structure capable of undergoing rapidly all kinds of vital transactions to respond promptly to different types of environmental cues, changes and stresses with focused and/or global reprogramming of gene expression. All this happens within one or a few ill-defined structures called “nucleoids” where the cellular DNA is localized.

The bacterial nucleoid is enclosed by the cytoplasm, likely separated from it by a physical chemical effect known as “molecular crowding” but not compartmentalized

by a nuclear envelope like that existing in eukaryotes. For many years, this circumstance, the size of the nucleoids, at the limits of resolution of the traditional detection methods of cell biology, and the elusiveness of their morphology and composition have made it particularly difficult to answer basic questions about the behavior and the structural and functional organization of the bacterial chromosome.

About 30 years ago, when I started being interested in the organization of the nucleoid and, more particularly, in the chemical nature, role and expression of the proteins associated with the bacterial chromosome, studies on this subject were at their infancy.

Indeed, a huge gap existed between the morphological information obtained through the pioneering studies of electron microscopists such as the late Professor Eduard Kellenberger and his colleagues and the almost non-existent biochemical characterization of the nucleoid and of its protein components. In 1977, Varshavsky had detected by SDS-PAGE the presence of two “histone-like” proteins within a purified *E. coli* deoxyribonucleoprotein preparation. He named the proteins B1 and B2 but, aside from their molecular weights, no other property was given, so that our present belief that these two proteins corresponded to HU and H-NS cannot be supported by any evidence. In fact, most scientists at that time considered the bacterial DNA to be “naked”, neutralized by mono- and divalent cations and polyamines and, given the absence of eukaryotic-type histones, they questioned the mere existence of DNA-associated architectural proteins.

Nuclease treatment of nucleoids obtained from gently lysed cells had already shown the existence of topologically independent domains of supercoiling as well as an “organizing” central core of RNA. While the latter turned out to be a preparation artifact, the existence of the topologically independent negatively supercoiled loops was later confirmed, initially by tri-methyl psoralen crosslinking and then by elegant site-specific recombination experiments and by accurate EM observations. The use of site-specific recombination between directly repeated *res* sites mediated by $\gamma\delta$ resolvase engineered to have different half-lives within the cell and the use of supercoiling sensitive reporter genes revealed the existence of approximately 450 domains of supercoiling per genome having a mean size of 11 kb and randomly located barriers. Further studies have also shown how the transcriptional activity of the chromosome may contribute to shaping the nucleoid and how rapidly disassembled nucleoid components can reassemble.

The separation of the chromosome into independent, negatively supercoiled loops, half of which are plectonemic, turns out to be of paramount importance not only as one of the mechanisms responsible for bacterial chromosome compaction within the nucleoid, but also for preventing the loss of DNA superhelicity. In fact, the existence of non-restrained negative supercoiling is required for a plethora of DNA functions and well known are the adverse, often lethal effects caused by both hyper- and hypo-supercoiling.

In addition to the aforementioned macro-molecular crowding and DNA supercoiling, an important role in DNA condensation is played by nucleoid-associated proteins which in the meantime have been identified and rigorously characterized. In fact, following a shaky and uncertain beginning which characterized the 1970s

and the first half of the 1980s, when several articles appeared reporting conflicting properties of ill-defined proteins supposedly associated with the chromosome and to which various names had been attributed, the major components of the nucleoid were finally thoroughly purified and their precise biochemical and genetic identities established. In this way, it was possible to discover that *E. coli* HU in reality consisted of two different polypeptide chains (HU α and HU β) whose amino acid sequences were promptly determined. Shortly thereafter, also the structural genes encoding these two proteins (*hupB* and *hupA*) were identified, mapped and sequenced and a close similarity between the two HU subunits and the two subunits of IHF (IHF-A and IHF-B) was detected. Likewise, the amino acid sequence of H-NS and the nucleotide sequence of its structural gene *hns* were determined. In turn, these data led to the detection of a close similarity between H-NS and StpA, a less abundant, yet probably not less important, nucleic acid binding protein. In the same period, two additional proteins (FIS and Lrp), which later turned out to be important components of the nucleoid, were also isolated and characterized.

It is now well established that these proteins are nucleoid structuring proteins which bind curved DNA, recognizing short, more or less degenerate consensus sequences, bend DNA and influence DNA supercoiling. In addition to contributing, through different mechanisms, to DNA compaction, at least some of these proteins participate in forming the dynamic barriers separating the topologically independent domains of supercoiling. Furthermore, it is also clear that the NAPs, in addition to being architectural proteins of the nucleoid, play other roles in the cell. In fact, several lines of evidence, including the highly pleiotropic effects displayed by mutations in their structural genes, indicate that the NAPs participate in DNA transactions such as recombination, repair and replication. Of particular relevance, in this connection, is the fact that all the NAPs, alone or in combination through synergistic or antagonistic mechanisms, have profound effects on the transcriptional activity of the cell.

The level of expression of the genes encoding NAPs is not constant during the growth cycle so that the intracellular concentration of these proteins varies as a function of the metabolic state of the cell and/or as a consequence of environmental changes. Since several promoters have been found to possess multiple, sometimes partially overlapping binding sites for these proteins, it is possible to envisage the existence of an intricate pattern of cross talks between the NAPs (e.g. the antagonistic effects of H-NS and FIS and HU and H-NS on the activity of some promoters) and the cyclic establishment or loss of integrated regulatory networks controlling global responses to environmental changes.

Taken together, all the data accumulated so far underlie the tight link existing between nucleoid architecture and nucleoid function and the close relationship between two apparently conflicting needs, namely that of condensing DNA and that of ensuring its accessibility through dynamic movements of the nucleoid and of its components.

Recent years have witnessed the development of new, powerful techniques to investigate the structure and functional organization of the bacterial nucleoid which have led to a renewed flourishing of the studies on this subject. Aside from the

aforementioned site-specific recombination, new microscopic techniques (e.g. confocal microscopy and AFM) and the manipulation of single and dual DNA molecules have contributed to giving a sharper image of the mechanisms by which the bacterial chromosome is condensed, made accessible and segregated. The picture that emerges is that of an analogic “machine” for which the most appropriate definition would be that of deterministic and organized chaos.

After studying the various chapters of this book, written by excellent scientists working at the forefront of this important aspect of molecular microbiology, the reader will certainly appreciate how much light has been shed on the bacterial nucleoid since the time it was considered stochastic chaos and bacterial DNA was regarded as “naked”. However, aside from realizing the extent of progress made in the last few years in understanding the nucleoid, the attentive reader will also perceive how much more remains to be learned.

Claudio O. Gualerzi

Contents

Part I Structure and Organization of the Bacterial Chromosome

1 Ultrastructure and Organization of Bacterial Chromosomes	3
Remus T. Dame	
2 Imaging the Bacterial Nucleoid	13
William Margolin	
3 The Chromosome Segregation Machinery in Bacteria	31
Peter L. Graumann	
4 Extrachromosomal Components of the Nucleoid: Recent Developments in Deciphering the Molecular Basis of Plasmid Segregation	49
Finbarr Hayes and Daniela Barillà	
5 Nucleoid Structure and Segregation	71
Conrad L. Woldringh	
6 Polymer Physics for Understanding Bacterial Chromosomes	97
Suckjoon Jun	
7 Molecular Structure and Dynamics of Bacterial Nucleoids.....	117
N. Patrick Higgins, B.M. Booker, and Dipankar Manna	
8 Nucleoid-Associated Proteins: Structural Properties	149
Ümit Pul and Rolf Wagner	
9 Dps and Bacterial Chromatin	175
Hanne Ingmer	

Part II Chromatin Organization in Archaea and Eukaryotes

- 10 Archaeal Chromatin Organization.....** 205
Stephen D. Bell and Malcolm F. White
- 11 The Topology and Organization of Eukaryotic Chromatin** 219
Andrew Travers and Georgi Muskhelishvili

Part III Regulation by Nucleoid-Associated Proteins

- 12 Bacterial Chromatin and Gene Regulation** 245
Charles J. Dorman
- 13 H-NS as a Defence System.....** 251
William Wiley Navarre
- 14 FIS and Nucleoid Dynamics upon Exit from Lag Phase** 323
Georgi Muskhelishvili and Andrew Travers
- 15 LRP: A Nucleoid-Associated Protein with Gene
Regulatory Functions.....** 353
Stacey N. Peterson and Norbert O. Reich
- 16 Extreme DNA Bending: Molecular Basis of the Regulatory
Breadth of IHF** 365
Amalia Muñoz, Marc Valls, and Víctor de Lorenzo
- 17 Role of HU in Regulation of *gal* Promoters** 395
Dale E.A. Lewis, Sang Jun Lee, and Sankar Adhya
- 18 Transcriptional Regulation by Nucleoid-Associated
Proteins at Complex Promoters in *Escherichia coli*.....** 419
Douglas F. Browning, David C. Grainger, Meng Xu,
and Stephen J.W. Busby
- Index.....** 445

Part I
Structure and Organization of the
Bacterial Chromosome

Chapter 1

Ultrastructure and Organization of Bacterial Chromosomes

Remus T. Dame

1.1 Introduction

Ever since the early observations by the Dutch microscopist Antonie van Leeuwenhoek in the late seventeenth century (communicated in a series of letters published in the Philosophical Transactions of the Royal Society) researchers have been fascinated by the ability to magnify and visualize cells and microorganisms microscopically. In eukaryotic cells due to their relatively large size and the separation from the cytoplasm by a membrane, cellular organelles such as the nucleus or mitochondria can be readily visualized in a simple light microscope. The situation is more complex in organisms that are several orders of magnitude smaller and in which the genetic material is not membrane-enclosed, such as bacteria and archaea. While by the end of the nineteenth century the nucleus and its mitotic dynamics had been resolved and the terms ‘chromatin’ and ‘chromosome’ had been coined, knowledge of a possible bacterial equivalent was still lacking. This was likely due to the fact that the chromosomal DNA of bacteria is translucent and featureless in the light microscope when not stained, and that the histological stains of that time (successfully applied to the nuclei of eukaryotic cells) were not successful in revealing the morphology of the genomic material of bacteria (Robinow and Kellenberger 1994). Despite the fact that a consistent morphology of the folded bacterial genome could not be described, bacterial cytologists during the first decades of the twentieth century became convinced that bacteria indeed contain ‘chromatin bodies’ (Delaporte 1939–1940). Particularly important for this development was the introduction of the Feulgen procedure and the Giemsa stain that specifically stain DNA and yielded ‘nucleoids’ of reproducible, regular morphology

R.T. Dame (✉)

Faculty of Mathematics and Natural Sciences, Leiden Institute of Chemistry
Laboratory of Molecular Genetics,

Einsteinweg 55, 2333 CC, Leiden, Netherlands;

Faculty of Science Division of Physics and Astronomy Section Physics of Complex Systems,
VU University Amsterdam, De Boelelaan 1081, 1081 HV, Amsterdam, The Netherlands

e-mail: rtdame@chem.leidenuniv.nl

(Neumann 1941; Piekarski 1937). The next big step forward in terms of resolving the nucleoid in much more detail was expected when electron microscopes became widely available in the late 1940s. Structures similar to the nucleoids observed in light microscopy studies (Piekarski 1937; Stempen 1950) were found, but, unlike the nuclei in eukaryotic cells, these had low electron density and did not resolve much additional detail (Hillier et al. 1949).

1.2 Global Structure of the Nucleoid: A Top-Down View

Despite a lot of effort by different investigators it seems that the general overall picture of the nucleoid (having an oval shape) remained unchanged for the next 50 years. Still, it was noted that the fine structure within the nucleoid depends on the fixation procedure used and the relevance of these structures is therefore unclear (Robinow and Kellenberger 1994). In the early 1990s a refinement of earlier models was proposed in which the nucleoid has a coralline structure with large excrescences extending from the nucleoid body (Bohrmann et al. 1991). This model is compatible with the notion that parts of the genome are attached to the membrane (Dworsky and Schaechter 1973). However, conventional light microscopy techniques have insufficient resolution to visualize the proposed excrescences and therefore this observation still awaits confirmation in the live cell. Fixation by rapid freezing is believed to conserve cells in a near-to-native state. This approach is used in conventional and novel electron microscopy techniques such as cryo electron tomography. The latter method has been effectively applied to imaging *E. coli* cells yielding beautiful images of the liquid-crystalline state of nucleoids in the stationary phase of growth (Frenkiel-Krispin et al. 2001; Wolf et al. 1999).

Another powerful approach taken up early on was to try to investigate isolated chromatin bodies or parts thereof. This had proven to be very informative in the case of eukaryotic chromatin, where it revealed the existence of chromatin filaments and also revealed fine-structure within these filaments in the form of nucleosomes clearly visible as ‘beads on a string’ (Finch et al. 1975; Olins and Olins 1974; Ris and Kubai 1970). These successes in the eukaryotic field were paralleled by only limited success in defining more accurately the structure of bacterial chromatin. The rosette-like structures of bacterial chromosomes as visualized by Kavenoff and Bowen (Kavenoff and Bowen 1976) are engraved into the minds of many generations of biologists. In fact one of their images turned into a “commercial icon” depicted on postcards, T-shirts etcetera. It is said that in that period one of the people in the field skipped the introductory slide in his lectures, referring to his T-shirt instead. In the studies of Kavenoff and Bowen *E. coli* cells were lysed in situ on an electron microscopy grid (in order to preserve as much as possible the integrity of the nucleoid) and directly prepared for imaging. These images have undoubtedly fed the imagination and sparked the interest of many a scientist. They are at the basis of the still current ideas about the organization of the bacterial chromosome in large topologically independent looped domains, that

found support in elegant *in vivo* recombination assays (Deng et al. 2005; Higgins et al. 1996) and global genomic approaches (Noom et al. 2007; Postow et al. 2004). Unlike in the case of their eukaryotic counterparts no proteins are found bound in these images of bacterial chromosomal DNA, unless they are first treated with cross-linking agents (Griffith 1976), which likely reflects these proteins being transiently bound to poorly defined positions along the DNA. After the detection of ‘nucleosome-like structures’ on DNA reconstituted with one of the most abundant proteins associated with the nucleoid, HU (Rouvière-Yaniv et al. 1979), the incorrect view that these proteins act like eukaryotic histones dominated the field for over two decades (Dame and Goosen 2002; van Noort et al. 2004). In retrospect, one can attribute the long persistence of this view, as well as the limited understanding of the action of the other architectural proteins associated with the nucleoid, largely to a lack of appropriate methodology.

One of the current methods that is less susceptible to the generation of artifacts, at least when compared to the electron microscopy protocols from the 1970s and 1980s, is scanning force microscopy. This technique relies on the construction of a topographic map of the object under investigation by scanning it with a nanometer-sized tip and is thus not suited to visualizing the nucleoid in the context of the intact cell. However, analogous to the approach employed by Kavenoff and Bowen, cells can be lysed on a surface and directly visualized. The images of nucleoids generated by this approach exhibit an interesting diversity of fibres of different diameters, proposed to reflect different orders of organization of bacterial chromatin (Kim et al. 2004) (proposed to be analogous to the various orders of organization observed in eukaryotic organisms). The images are even qualitatively different depending on the levels of expression of proteins believed to be involved in nucleoid organization (see below) (Ohniwa et al. 2006). However, as the cell is rich in proteins and RNA, and as with gentle lysis fragments of the peptidoglycan layer may remain, it is also here not clear to what extent these images are a true reflection of the *in vivo* situation.

The method of choice to visualize nucleoid structure and dynamics *in vivo* currently is widefield epifluorescence or confocal microscopy employing direct chemical staining of DNA (using fluorescent intercalating dyes) or fluorescently tagged fusion proteins localizing to the nucleoid (see the contributions of William Margolin, Conrad Woldringh and Finnbar Hayes and Daniela Barillà). Both approaches have potential drawbacks: the intercalating dyes may affect DNA structure and compete with DNA binding proteins, while the fluorescent tags fused to the proteins may affect their binding properties. However, this does not seem to affect the overall low-resolution picture of the nucleoid as currently obtained with these methods. The emphasis has to date been on the mere staining of the nucleoid, often to provide a reference for the localization of other fluorescently tagged proteins (Giangrossi et al. 2001; Wery et al. 2001). In a different approach, specific sites along the genome are labeled (for instance, using LacI-GFP targeting *lac* operator sites inserted at a defined position) rather than the nucleoid as a whole. This approach has proven to be particularly powerful in studies on the (directed) movement of individual loci in the nucleoid (for instance, during chromosome segregation), in correlating the physical position on the nucleoid with the linear

location along the genome and in determining the local ‘fluidity’ of the nucleoid (from the freedom of movement of these sites) (Elmore et al. 2005; Teleman et al. 1998; Viollier et al. 2004) (see the contributions of Conrad Woldringh and Suckjoon Jun).

The advantages of confocal microscopy in terms of imaging quality are only limited in small microbes, but it has brought into reach approaches that can reveal the binding dynamics of proteins *in vivo* such as FRAP (Fluorescent Recovery After Photobleaching) and FLIP (Fluorescence Loss In Photobleaching) (Mullineaux 2007). Whereas these approaches are in common use for studies on the eukaryotic nucleus (Koster et al. 2005; van Royen et al. 2009), investigators have been hesitant in applying them to bacterial cells, likely as the diameter of the nucleoid is of the same order as that of the diffraction limited laser spot.

Accurate segregation of chromosomes as well as plasmids is of vital importance for the cell to ensure proper transfer of its genetic information to the daughter cells during cell division. The segregation process and subsequent positioning of the chromosome within the daughter cells has been widely studied by microscopy employing fluorescent labels at the origin, terminus and intermediate positions. A number of distinct non-mutually exclusive mechanisms seems to be employed for segregation of chromosomes and plasmids. Plasmids generally require an active mechanism of segregation that involves cytoskeletal components. Such active mechanisms have a large appeal to investigators, due to their analogy to the known mechanisms operating in eukaryotes. However, there is increasing evidence that segregation of some plasmids as well as chromosomes occurs by “passive mechanisms” and that chromosome segregation in bacteria also does not need an active mechanism (see the contributions of Peter Graumann, Finnbar Hayes and Daniela Barillà, Suckjoon Jun and Conrad Woldringh). In fact, it has recently been suggested that chromosome segregation and ordering can be explained largely based on entropic considerations (Jun and Mulder 2006).

1.3 Mechanisms of Local and Global Nucleoid Organization: A Bottom-Up View

Besides a top down approach where the nucleoid is investigated with its overall *in vivo* structure as a starting point, a lot of studies have aimed at characterizing the individual (molecular) players in nucleoid organization. Three key factors with such a role have been identified: architectural proteins, DNA supercoiling and macromolecular crowding. The architectural proteins of bacteria and archaea (as described in the contributions of Pul and Wagner and Bell and White respectively) are generally found associated with the nucleoid (Azam et al. 2000; Varshavsky et al. 1977) and therefore believed to play a central role in nucleoid organization. These proteins do not exhibit sequence or structural conservation across the three kingdoms of life, but their architectural modes of action on the genome appear very similar (Luijsterburg et al. 2008; Oberto et al. 1994). The possible parallels in terms of higher-order genome organization in eukaryotes and bacteria are discussed

in Chapter 11. Besides a generic role in overall organization of the genome, these proteins are involved as co-factors in an ever-expanding repertoire of DNA transactions. As such they also play important roles in (global) regulation of transcription, as discussed in Chapters 12–18. DNA supercoiling is the folding of DNA into higher order structures due to torsion in the DNA duplex. This results in reduction of the effective volume of the DNA. Supercoiling is maintained by the action of a family of specialized enzymes: topoisomerases (Drlica 1992; Wang 1985) and is discussed in the contributions of Pat Higgins and colleagues and Peter Graumann). Macromolecular crowding is a physico-chemical contribution to DNA compaction that derives from the high concentration of macromolecules such as RNA and proteins, which promotes a phase separation between DNA and cytoplasm (Odijk 1998; Zimmerman and Murphy 1996) (see the contributions of Conrad Woldringh and Suckjoon Jun). A major challenge is to identify the individual contributions of these three factors. To date this is still largely unresolved as they are tightly interconnected and likely act cooperatively (Luijsterburg et al. 2008). The most accessible to investigations in a reductionist system are the nucleoid-associated proteins. Since their first identification in the 1970s (Rouvière-Yaniv and Gros 1975; Spassky and Buc 1977; Varshavsky et al. 1977) these proteins have been extensively biochemically and biophysically characterized (Dame 2005). A bulk of structural and functional information is therefore available to date. However, due to the redundant function in genome organization of many NAP's, as well as their activity being so tightly interconnected with supercoiling and macromolecular crowding the exact function of these proteins in genome organization is hard to assign *in vivo*.

1.4 Integrating the Top-Down and Bottom-Up Approach

Our understanding of many aspects of bacterial chromatin organization has increased a lot over the last few decades. However, there is still a large gap between the global view of the nucleoid and the role of the individual factors involved in imposing this structure onto the genome. In particular, the dynamics of nucleoid organization are still poorly understood. These are likely very important as it is known that expression levels of nucleoid-associated proteins are strongly affected by environmental stimuli and that nucleoid-associated proteins are key components in the adaptive response of the cell (Dorman 2009; Hengge-Aronis 1999). There is a delicate interplay between different nucleoid-associated proteins at the level of transcription regulation (see the contribution of Steve Busby and colleagues) and similar interplay likely occurs on the global scale where these proteins act antagonistically or cooperatively to set the local compaction state, as well as to modulate the overall degree of compaction (Dame 2005; Luijsterburg et al. 2006; Maurer et al. 2009; Travers and Muskhelishvili 2005). This begs for an integrated approach aimed at correlating the genomic location of NAP's with local DNA density and transcriptional activity. Several techniques are already available and have in part

been applied to addressing these types of questions. Genome-wide localization studies based on Chromatin Immuno Precipitation are becoming mainstream and currently benefit in terms of time input and resolution from high-throughput deep sequencing methods (Bulyk 2006; Robertson et al. 2007; Wade et al. 2007). Coupled to analysis of global gene expression patterns such studies can provide direct indications of regulatory roles for nucleoid-associated proteins. In parallel, as in vivo imaging methods gain higher resolution and sensitivity, individual fluorescently tagged proteins can now be localized and followed over time in the live cell with high accuracy (Xie et al. 2008). This type of approach in combination with genomic labels at defined positions allows spatial mapping of the localization of such proteins. Important in this regard is the development of so called ‘super resolution imaging techniques’, which through ingenious engineering solutions facilitate optical resolutions far below the diffraction limit (Gitai 2009; Hell 2007; Hess et al. 2006; Rust et al. 2006). A drawback of these approaches is that they still require labeling of DNA or protein, which may lead to system perturbations. In that light potentially less-invasive methods such as those emerging in other fields of science appear very attractive. For instance, in-cell NMR (Nuclear Magnetic Resonance) spectroscopy (Augustus et al. 2009; Charlton and Pielak 2006; Sakakibara et al. 2009) or label-free optical techniques (Fujita and Smith 2008) may in the near future evolve into useful complements of the more conventional approaches.

1.5 Conclusion

Simple as they may superficially appear, bacteria and the organization of their genomes is still far from being understood. Such knowledge is obviously crucial for understanding bacterial physiology and the interplay between bacteria and their environment. Interesting and important in their own right bacteria also act as an accessible system to reveal, explore and quantitatively describe the principles of genome organization applying to all forms of life. To date even a ‘complete’ description of three-dimensional genome organization and dynamics in bacteria (combining the knowledge on the action of architectural proteins, with their genome-wide and spatial localization and fundamental physical principles) no longer seems remote.

Acknowledgements Research in the author’s laboratory is supported by grants from the Netherlands Organization for Scientific Research (NWO).

References

- Augustus AM, Reardon PN, Spicer LD (2009) MetJ repressor interactions with DNA probed by in-cell NMR. *Proc Natl Acad Sci USA* 106:5065–5069
- Azam TA, Hiraga S, Ishihama A (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* 5:613–626

- Bohrmann B, Villiger W, Johansen R, Kellenberger E (1991) Coralline shape of the bacterial nucleoid after cryofixation. *J Bacteriol* 173:3149–3158
- Bulyk ML (2006) DNA microarray technologies for measuring protein-DNA interactions. *Curr Opin Biotechnol* 17:422–430
- Charlton LM, Pielak GJ (2006) Peeking into living eukaryotic cells with high-resolution NMR. *Proc Natl Acad Sci USA* 103:11817–11818
- Dame RT (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* 56:858–870
- Dame RT, Goosen N (2002) HU: promoting or counteracting DNA compaction? *FEBS Lett* 529:151–156
- Delaporte B (1939–1940) Recherches cytologiques sur les bactéries et les cyanophycées. *Rev Gén Bot* 51–52:615–643, 689–708, 748–768, 112–660
- Deng S, Stein RA, Higgins NP (2005) Organization of supercoil domains and their reorganization by transcription. *Mol Microbiol* 57:1511–1521
- Dorman CJ (2009) Global regulators and environmental adaptation in Gram-negative pathogens. *Clin Microbiol Infect* 15(Suppl 1):47–50
- Drlica K (1992) Control of bacterial DNA supercoiling. *Mol Microbiol* 6:425–433
- Dworsky P, Schaechter M (1973) Effect of rifampin on the structure and membrane attachment of the nucleoid of *Escherichia coli*. *J Bacteriol* 116:1364–1374
- Elmore S, Muller M, Vischer N, Odijk T, Woldringh CL (2005) Single-particle tracking of *oriC*-GFP fluorescent spots during chromosome segregation in *Escherichia coli*. *J Struct Biol* 151:275–287
- Finch JT, Noll M, Kornberg RD (1975) Electron microscopy of defined lengths of chromatin. *Proc Natl Acad Sci USA* 72:3320–3322
- Frenkel-Krispin D, Levin-Zaidman S, Shimoni E, Wolf SG, Wachtel EJ, Arad T, Finkel SE, Kolter R, Minsky A (2001) Regulated phase transitions of bacterial chromatin: a non-enzymatic pathway for generic DNA protection. *EMBO J* 20:1184–1191
- Fujita K, Smith NI (2008) Label-free molecular imaging of living cells. *Mol Cells* 26:530–535
- Giangrossi M, Exley RM, Le Hegarat F, Pon CL (2001) Different in vivo localization of the *Escherichia coli* proteins CspD and CspA. *FEMS Microbiol Lett* 202:171–176
- Gitai Z (2009) New fluorescence methods for microbiology: sharper, faster and quantitative. *Curr Opin Microbiol* 12:341–346
- Griffith JD (1976) Visualization of prokaryotic DNA in a regularly condensed chromatin-like fiber. *Proc Natl Acad Sci USA* 73:563–567
- Hell SW (2007) Far-field optical nanoscopy. *Science* 316:1153–1158
- Hengge-Aronis R (1999) Interplay of global regulators and cell physiology in the general stress response of *Escherichia coli*. *Curr Opin Microbiol* 2:148–152
- Hess ST, Girirajan TP, Mason MD (2006) Ultra-high resolution imaging by fluorescence photo-activation localization microscopy. *Biophys J* 91:4258–4272
- Higgins NP, Yang X, Fu Q, Roth JR (1996) Surveying a supercoil domain by using the gamma delta resolution system in *Salmonella typhimurium*. *J Bacteriol* 178:2825–2835
- Hillier J, Mudd S, Smith AG (1949) Internal structure and nuclei in cells of *Escherichia coli* as shown by improved electron microscopic techniques. *J Bacteriol* 57:319–338
- Jun S, Mulder B (2006) Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proc Natl Acad Sci USA* 103:12388–12393
- Kavenoff R, Bowen BC (1976) Electron microscopy of membrane-free folded chromosomes from *Escherichia coli*. *Chromosoma* 59:89–101
- Kim J, Yoshimura SH, Hizume K, Ohniwa RL, Ishihama A, Takeyasu K (2004) Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. *Nucleic Acids Res* 32:1982–1992
- Koster M, Frahm T, Hauser H (2005) Nucleocytoplasmic shuttling revealed by FRAP and FLIP technologies. *Curr Opin Biotechnol* 16:28–34
- Luijsterburg MS, Noom MC, Wuite GJ, Dame RT (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol* 156:262–272

- Luijsterburg MS, White MF, van Driel R, Dame RT (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol* 43:393–418
- Maurer S, Fritz J, Muskhelishvili G (2009) A systematic in vitro study of nucleoprotein complexes formed by bacterial nucleoid-associated proteins revealing novel types of DNA organization. *J Mol Biol* 387:1261–1276
- Mullineaux CW (2007) Localization and mobility of bacterial proteins by confocal microscopy and fluorescence recovery after photobleaching. *Methods Mol Biol* 390:3–15
- Neumann F (1941) Untersuchungen zur Erforschung der Kernverhältnisse bei den Bakterien. *Zentralbl Bakteriol Parasitenkd Infektionslr Hyg Abt I Orig* 103:385–400
- Noom MC, Navarre WW, Oshima T, Wuite GJ, Dame RT (2007) H-NS promotes looped domain formation in the bacterial chromosome. *Curr Biol* 17:R913–R914
- Oberto J, Drlica K, Rouvière-Yaniv J (1994) Histones, HMG, HU, IHF: Mème combat. *Biochimie* 76:901–908
- Odijk T (1998) Osmotic compaction of supercoiled DNA into a bacterial nucleoid. *Biophys Chem* 73:23–29
- Ohniwa RL, Morikawa K, Kim J, Ohta T, Ishihama A, Wada C, Takeyasu K (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. *EMBO J* 25:5591–5602
- Olins AL, Olins DE (1974) Spheroid chromatin units (v bodies). *Science* 183:330–332
- Piekarski G (1937) Zytologische Untersuchungen an *Paratyphus* und *Coli* Bakterien. *Arch Mikrobiol* 8:428–429
- Postow L, Hardy CD, Arsuaga J, Cozzarelli NR (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev* 18:1766–1779
- Ris H, Kubai DF (1970) Chromosome structure. *Annu Rev Genet* 4:263–294
- Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A, Thiessen N, Griffith OL, He A, Marra M, Snyder M, Jones S (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 4:651–657
- Robinow C, Kellenberger E (1994) The bacterial nucleoid revisited. *Microbiol Rev* 58:211–232
- Rouvière-Yaniv J, Gros F (1975) Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*. *Proc Natl Acad Sci USA* 72:3428–3432
- Rouvière-Yaniv J, Yaniv M, Germond JE (1979) *E. coli* DNA binding protein HU forms nucleosomelike structure with circular double-stranded DNA. *Cell* 17:265–274
- Rust MJ, Bates M, Zhuang X (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* 3:793–795
- Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Walchli M, Smith BO, Shirakawa M, Guntert P, Ito Y (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* 458:102–105
- Spassky A, Buc HC (1977) Physico-chemical properties of a DNA binding protein: *Escherichia coli* factor H1. *Eur J Biochem* 81:79–90
- Stempen H (1950) Demonstration of the chromatinic bodies of *Escherichia coli* and *Proteus vulgaris* with the aid of the phase contrast microscope. *J Bacteriol* 60:81–87
- Teleman AA, Graumann PL, Lin DC, Grossman AD, Losick R (1998) Chromosome arrangement within a bacterium. *Curr Biol* 8:1102–1109
- Travers A, Muskhelishvili G (2005) Bacterial chromatin. *Curr Opin Genet Dev* 15:507–514
- van Noort J, Verbrugge S, Goosen N, Dekker C, Dame RT (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc Natl Acad Sci USA* 101:6969–6974
- van Royen ME, Farla P, Mattern KA, Geverts B, Trapman J, Houtsmuller AB (2009) Fluorescence Recovery After Photobleaching (FRAP) to study nuclear protein dynamics in living cells. *Methods Mol Biol* 464:363–385
- Varshavsky AJ, Nedospasov SA, Bakayev VV, Bakayeva TG, Georgiev GP (1977) Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Res* 4:2725–2745

- Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Natl Acad Sci USA* 101:9257–9262
- Wade JT, Struhl K, Busby SJ, Grainger DC (2007) Genomic analysis of protein-DNA interactions in bacteria: insights into transcription and chromosome organization. *Mol Microbiol* 65:21–26
- Wang JC (1985) DNA topoisomerases. *Annu Rev Biochem* 54:665–697
- Wery M, Woldringh CL, Rouvière-Yaniv J (2001) HU-GFP and DAPI co-localize on the *Escherichia coli* nucleoid. *Biochimie* 83:193–200
- Wolf SG, Frenkiel D, Arad T, Finkel SE, Kolter R, Minsky A (1999) DNA protection by stress-induced biocrystallization. *Nature* 400:83–85
- Xie XS, Choi PJ, Li GW, Lee NK, Lia G (2008) Single-molecule approach to molecular biology in living bacterial cells. *Annu Rev Biophys* 37:417–444
- Zimmerman SB, Murphy LD (1996) Macromolecular crowding and the mandatory condensation of DNA in bacteria. *FEBS Lett* 390:245–248

Chapter 2

Imaging the Bacterial Nucleoid

William Margolin

Abstract This chapter outlines how important properties of the bacterial nucleoid have been discovered by direct visualization of the nucleoid in situ by microscopy. Relatively new tools for these investigations include fluorescent protein fusions, in situ hybridization, cryo-EM, and atomic force microscopy. The nucleoid is not only just a passive carrier of chromosomal DNA, but also actively influences global organization of the cell including placement of the division site to prevent unwanted cutting of the nucleoid by the division septum. Possibly because of this key role in cellular organization, nucleoids are positioned in specific locations in the cell, and certain mechanisms such as FtsK-mediated DNA transport keep DNA away from the division septum. Condensins and other nucleoid-associated proteins help to maintain nucleoids in a compacted state, in part to facilitate proper segregation to daughter cells. In addition, RNA and protein synthesis seem to act in a balance to maintain overall nucleoid shape. During cellular differentiation to and from dormant states, nucleoid shape and density can vary dramatically, probably reflecting the need to protect the DNA. Finally, microscopic imaging has just begun to elucidate the great diversity of nucleoid organization in bacterial species.

Keywords Cytoskeleton • bacteria • localization • fluorescence • GFP • nucleoid • actin • tubulin • protein • cytokinesis • segregation

2.1 A Brief History of Visualizing the Bacterial Nucleoid

The defining characteristic of prokaryotes is that their chromosomal DNA, unlike that of eukaryotes, is not enclosed in a membrane-bound nucleus. Despite this, bacterial chromosomal DNA remains organized in a defined structure called the nucleoid. First coined by Piekarski in the 1930s (Piekarski 1937), the nucleoid

W. Margolin (✉)

University of Texas Medical School-Houston, Houston, Texas, USA

e-mail: William.Margolin@uth.tmc.edu

forms a clearly distinct phase from the rest of the cytoplasm. Nucleoids were originally visualized by staining fixed cells with Faelgen and Giemsa dyes, although these methods often produced artifacts (Robinow and Kellenberger 1994). To minimize these artifacts, nucleoids were also visualized in living cells directly under phase contrast (Stempen 1950). Subsequently, finer morphological details of the nucleoid were uncovered by placing the cells in high concentrations of gelatin, which increased contrast by making the refractive index of the growth medium similar to that of the cell cytoplasm (Mason and Powelson 1956; Yamaichi and Niki 2004). Because of the limits of light microscopy, higher resolution could only be achieved by transmission electron microscopy (EM) of fixed cell sections that were dehydrated and embedded in resin, and later by freeze-substitution of unfixed samples. Unfortunately, this higher resolution also results in significant distortion of the native nucleoid morphology (Eltsov and Zuber 2006). One reason for this is that the protein density of the bacterial nucleoid is low compared to the histone-rich eukaryotic chromosome (Bendich and Drlica 2000), such that bacterial DNA tends to aggregate more readily during specimen preparation.

Light microscopic studies consistently showed that nucleoids of growing bacteria such as *Escherichia coli* occupy a significant portion of the cytoplasmic space, are rather irregular in shape despite being clearly coalesced into a single mass, and duplicate prior to cell division (Fig. 2.1). Higher resolution EM studies suggested that this

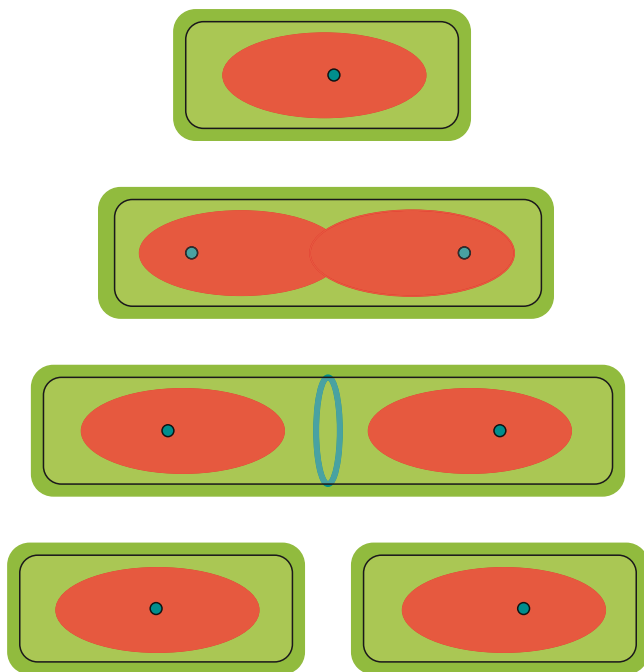


Fig. 2.1 The nucleoid during the cell division cycle of *E. coli*. Shown is a diagram of an *E. coli* cell at several stages during a division cycle, with the nucleoid in red, the Z ring in cyan, and *oriC* as a blue circle

irregular shape results from many projections of the nucleoid mass into the cytoplasm to form a “coralline” shape (Bohrmann et al. 1991). During the early stages of duplication, the nucleoid is sometimes observed as a bilobed shape (Yamaichi and Niki 2004; Zimmerman 2003). Before cell division occurs, the two lobes separate into two distinct nucleoids. In rod-shaped bacteria such as *E. coli* or *Bacillus subtilis*, the nucleoids generally separate by partitioning along the cell’s long axis, ensuring that each daughter will receive one nucleoid after binary fission.

Other useful tools have been developed to visualize nucleoids. Improved DNA-specific fluorescent dyes such as DAPI, Hoechst, and SYTO stains can illuminate nucleoid shape and dynamics in living or fixed cells (Fig. 2.2), and their high sensitivity can be used to confirm loss of the nucleoid under certain conditions (see below). DAPI, which emits in the blue range, is particularly useful in conjunction with other fluorophores such as GFP, membrane stains such as FM4–64, or immunostaining techniques to simultaneously visualize DNA, protein, and membrane localization and dynamics.

Even with all these tools, the basic shape and dynamics of the whole nucleoid under the light microscope look about the same now as they did 50 years ago. However, several relatively recent breakthroughs in imaging have shed new light on the organization and dynamics of the chromosomal DNA within the nucleoid. One of these is the ability to monitor the location of specific segments of the intact chromosome. Fluorescence in situ hybridization, or FISH, can label any genetic locus with a fluorescent DNA probe specific for that DNA sequence. Originally developed for eukaryotic chromosomes, it was adapted for bacteria about 10 years ago (Niki and Hiraga 1998). Because of the wide spectrum of fluorophores available for conjugating to DNA probes, the location of multiple loci can be visualized simultaneously without the need for genetic modification of the strains. A disadvantage of FISH is that live cells cannot be used because of the need for membrane permeabilization. However, cells can either be grouped by size as a proxy for cell age, or synchronized prior to fixation, to obtain time-dependent profiles of chromo-

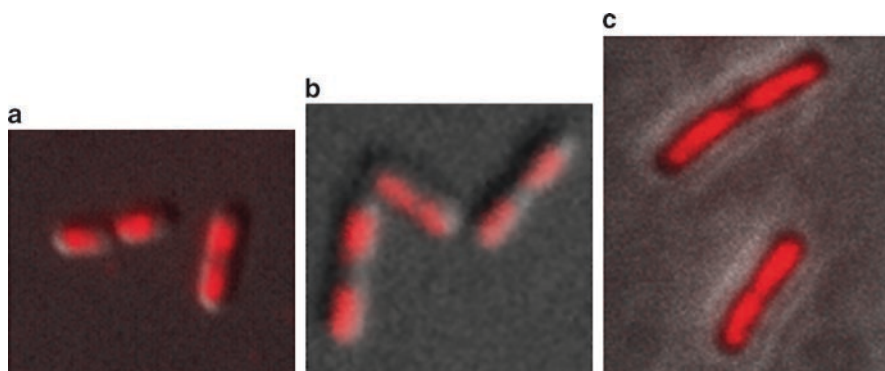


Fig. 2.2 Imaging the nucleoid by light microscopy with fluorescent stains. Logarithmically-growing *E. coli* cells were incubated with SYTO 16 (a) or DAPI (b) and the live cells were imaged by fluorescence microscopy. *E. coli* was also grown similarly in the absence of any stain, fixed with methanol, and subsequently stained with DAPI (c). For SYTO and DAPI, the blue or green emission light was pseudocoloured red for greater contrast

some dynamics. FISH was used to demonstrate that bacterial chromatin is organized spatially in parallel with chromosomal gene position (Niki et al. 2000).

For monitoring the localization and dynamics of specific chromosomal loci in live cells, the chromosome is first engineered to carry one or more tandem arrays of a high-affinity binding site such as the *lac* or *tet* operator (Webb et al. 1997). The insertion of one of these arrays at a specific chromosomal locus allows this segment of DNA to be localized in real time, because these strains are also engineered to express a Lac or Tet repressor protein genetically fused to a fluorescent protein such as green fluorescent protein (GFP) or a red fluorescent protein such as mCherry. Binding of a fluorescently labeled repressor to its cognate array of DNA binding sites results in a fluorescent focus inside the cell that is visible with fluorescence microscopy (Fig. 2.3). If the Tet repressor protein is labeled with GFP and the Lac repressor is labeled with mCherry, for instance, then two chromosomal loci can be monitored simultaneously in time-lapse movies. These methods were used to show that the replication origin (*oriC*) and terminus reside at opposite ends of the nucleoid, with the intermediate chromosomal loci positioned in sequence between them (Teleman et al. 1998; Viollier et al. 2004).

Moreover, other proteins that bind naturally to sites in the chromosome can be fluorescently labeled without the need to engineer a special binding site array, provided the sensitivity of detection is sufficiently high. For example, ParB proteins bind to centromere-like sites on the chromosome close to *oriC*, and thus serve as markers for the location of *oriC* at any time throughout the cell cycle (Thanbichler and Shapiro 2006). Similarly, the SeqA protein also binds near *oriC*, helping to keep the replication origin sequestered between firings (Fig. 2.3). These new cytological methods have paved the way for important insights into how chromosomes are organized within the nucleoid and will be elaborated in later chapters.

Another major technical breakthrough in imaging is the development of cryo-electron microscopy of vitreous sections. This method can visualize cytoplasmic contents of bacteria in their native hydrated state, without artifacts from chemical fixatives or freeze-substitution. Regions of high contrast, such as cell membranes, can be observed with unprecedented clarity. However, contrast for other regions of the cell is often quite low, and the nucleoid, while visible, is often difficult to distinguish from the rest of the cytoplasm in many cryo-EM images (Eltsov and Dubochet 2005; Eltsov and Zuber 2006). Further advances in 3-D tomographic reconstruction should solve this problem, and may help to elucidate fine structural details of the intact nucleoid that so far have been elusive.

Finally, atomic force microscopy (AFM) has been used recently to examine the structures of different chromosomal subdomains. Because AFM measures surface topography, nucleoids must be released from cells by lysis in situ (Ohniwa et al. 2006). However, the high resolution and contrast of AFM can distinguish among DNA fibers of different widths in the 10–100 nm range, which is useful for describing the fine structure of nucleoid domains in cells under various growth conditions. Similar lysis in situ was used previously with lower-resolution fluorescence techniques to determine structural differences between nucleoids of different species (Hinnebusch and Bendich 1997).

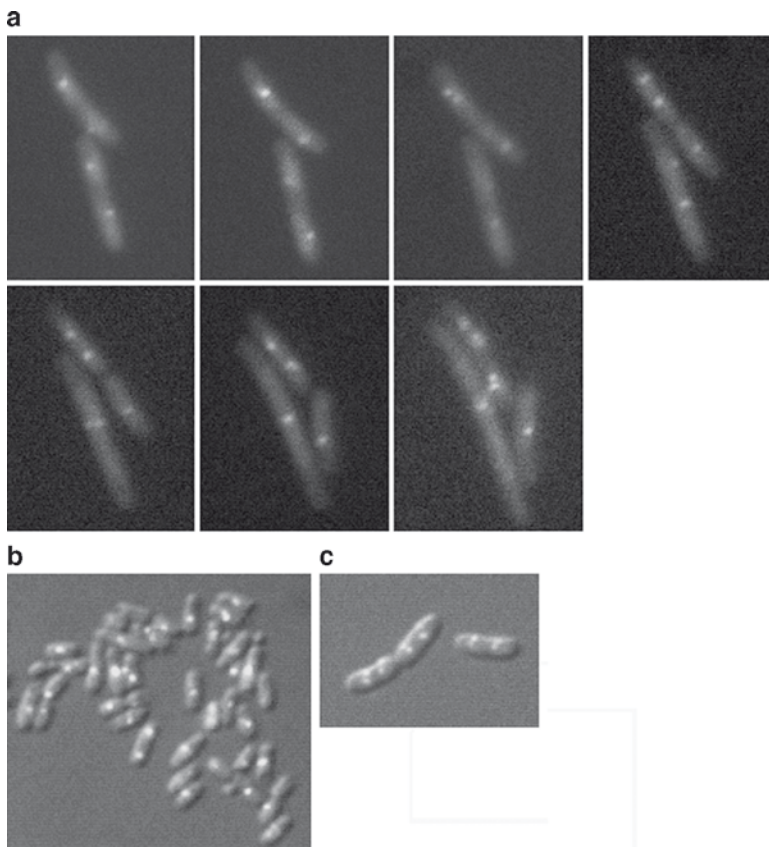


Fig. 2.3 Tracking the position of *oriC* within the *E. coli* nucleoid. **(a)** Cells of an *E. coli* strain (WM1075) containing a tandem *lacO* array inserted near *oriC* and expressing a GFP fusion to lac repressor (LacI) from a plasmid. The cells were grown in minimal glucose medium on an agarose pad and imaged with fluorescence microscopy approximately every 10 min over the course of 70 min (top left to bottom right). The fluorescent foci represent the sites at which multiple LacI-GFP molecules bind to the *lacO* array, reflecting the position of *oriC*. Duplication of one *oriC* is visible in the upper cell at the third time point, and significant separation of the duplicated *oriC*s is visible in the fourth time point prior to cell separation. **(b–c)** Shown are cells expressing a *seqA-gfp* fusion grown in minimal medium **(b)** or rich medium **(c)**. Fast growing cells usually have two SeqA foci, in contrast to slow growing cells, which usually have only one focus

2.1.1 The Nucleoid and the Cell Cycle

To ensure that progeny cells receive a copy of the genome, the nucleoid must be duplicated and properly positioned for cell division or cell budding. In bacteria that divide by binary fission such as *E. coli* or *B. subtilis*, the nucleoid is observed as a single mass of DNA in newborn cells. During fast growth, these newborn cell nucleoids will already contain actively replicating chromosomes. As a result, most

fast growing cells will have nucleoids in the process of separating. Slow growing cells, on the other hand, start chromosomal replication after a G1-like period. Therefore, their newborn cells will have a single nucleoid with no sign of separation, and a higher percentage of slow growing cells versus fast growing cells will have a single nucleoid (Woldringh 1976). As the cell cycle progresses, this single nucleoid will gradually assume a bilobed appearance, indicative of replication and ongoing chromosome partitioning into the future daughter cells. In either fast or slow growing cells, these nucleoids will need to be placed on either side of the cell centre, away from the cytokinesis machinery. Some components of nucleoid transport will be briefly overviewed below, but a much more detailed discussion appears in a later chapter devoted entirely to chromosome segregation (Chapter 3)

The first known component of cytokinesis is the Z ring, composed of assembled filaments of the FtsZ protein (Bi and Lutkenhaus 1991). Under most conditions, the Z ring forms at the centre of the cell only when the two daughter nucleoids are visibly separated on either side (Fig. 2.4). This is important, because if the Z ring assembled at the cell centre over an unpartitioned nucleoid, cytokinesis could guillotine the nucleoid, causing destruction of the chromosome and cell death. Although the temporal regulation of Z ring assembly is not understood, the spatial regulation of Z ring assembly such that it usually does not assemble over a nucleoid is well supported. The original evidence came from studies of division septa, showing that they can appear in most locations in the cell except on top of nucleoids (Woldringh et al. 1990). This “nucleoid occlusion” effect was also observed with Z rings, suggesting that FtsZ is the target of the inhibition (Yu and Margolin 1999). Despite the strong preference of the Z ring for the cell centre, Z rings usually form on one side of the nucleoid, away from the cell centre, if chromosome replication is inhibited (Harry et al. 1999; Sun and Margolin 2001) (Fig. 2.4). If replication proceeds but partitioning is inhibited, as is the case in mutants defective in Topoisomerase IV such as *parC(ts)*, then Z rings will form throughout the cell except on top of nucleoids. In *E. coli*, this effect is most dramatic when cell division is inhibited, as the cells form long filaments that permit iteration of the localization pattern (Yu and Margolin 1999) (Fig. 2.4).

The molecular mechanism of nucleoid occlusion is not completely understood, but has strong support from the discovery of specific inhibitor proteins, called Noc and SlmA in *B. subtilis* and *E. coli*, respectively (Bernhardt and de Boer 2005; Wu and Errington 2004). These bifunctional proteins bind DNA and localize throughout the nucleoid. In addition, they help to inhibit Z ring assembly over the nucleoid. When Noc or SlmA are inactivated along with the Min system, another key negative regulator of Z ring assembly, Z rings assemble promiscuously and often form on top of the nucleoid. If chromosome replication is blocked, many of these rings will trigger cytokinesis and guillotine the nucleoid, resulting in nucleoid-free cells (see below). Therefore, in *B. subtilis* and *E. coli*, the nucleoid not only keeps the chromosome spatially organized, but also regulates global cellular organization via these and possibly other nucleoid-occlusion proteins to ensure that the chromosome is stably inherited in every daughter cell. However, other species, such as *Streptococci* and *Corynebacteria*, form Z rings directly over nucleoids that have not yet significantly partitioned (Morlot et al. 2003; Ramos et al. 2005). Therefore,

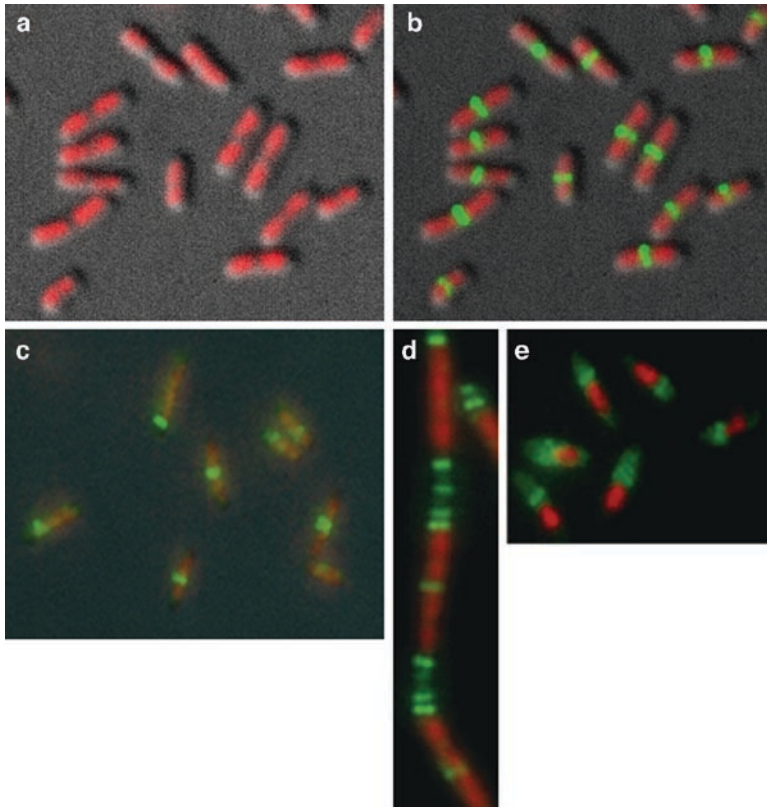


Fig. 2.4 Z ring localization with respect to nucleoids. Cells of either wild-type (**a–b**), a *dnaA(ts)* mutant at the nonpermissive temperature (**c**) a *parC(ts) Δmin ftsA(ts)* mutant at the nonpermissive temperature (**d**) or a *gyrB(ts)* mutant at the nonpermissive temperature (**e**) were fixed, stained with DAPI (pseudocolored red) and immunostained for FtsZ (green) except for panel A, which lacks the FtsZ staining to highlight the space between separated nucleoids at midcell. Under these conditions, the *dnaA(ts)* mutant fails to initiate DNA replication, the *gyrB(ts)* mutant is defective in negative DNA supercoiling, and the *parC(ts) Δmin ftsA(ts)* mutant simultaneously lacks topoisomerase IV (required for proper decatenation and partitioning of chromosomes), the Min system, and the FtsA protein required for cytokinesis, resulting in nondividing filaments with large regions devoid of nucleoids that assemble Z rings promiscuously. Panels **b–c** are adapted from (Sun and Margolin 2001) Copyright© American Society for Microbiology; Panel D is adapted from Yu and Margolin (1999)

post-septational transport of the nucleoid away from the division septum by FtsK/ SpoIIIE may be used instead of nucleoid occlusion in these species (see below).

Nucleoid-free cells also arise when the other system that positions the Z ring, the Min system, is inactivated. The Min proteins form a morphogenic gradient throughout the length of the cell, inhibiting Z ring formation at the cell poles and stimulating Z ring formation at the cell centre, between the two separate daughter nucleoids (Rothfield et al. 2005). In cells lacking the Min proteins, nucleoid occlusion becomes the dominant spatial regulator for the Z ring. As a result, cells divide either at the

nucleoid-free zone near a cell pole or at midcell after nucleoid partitioning. Cytokinesis at the poles generates nucleoid-free minicells that are usually spheroidal, with only polar cell wall material (Adler et al. 1967). Their lack of DNA is more striking when minicells are produced in chains of cells that have a separation defect (Yu et al. 1998b) (Fig. 2.5). Although minicells are ultimately inviable because they lack a chromosome, they remain metabolically active for a considerable time and have been used to produce radiochemically pure protein expressed from plasmids. They are therefore a good system to study cell functions in the absence of a nucleoid.

2.1.2 Factors that Position the Nucleoid in the Cell

The nucleoids of rod-shaped cells such as *E. coli* or *B. subtilis* are consistently positioned at the cell centre, leaving nucleoid-free areas at each cell pole. One of the reasons why minicell-producing strains are viable is because cell division at the poles pinches off nucleoid-free minicells without perturbing the central nucleoid. The mechanism regulating this positioning is not understood, but there is much evidence indicating that some dedicated nucleoid positioning mechanism must exist. For example, after a block to cytokinesis, cells continue to elongate, replicate their chromosomes, and separate nucleoids from each other. This separation cannot involve tethering to a cell pole, because these filamentous cells can become tens of microns long, with multiple nucleoids, and remain viable for a time (e.g., they can

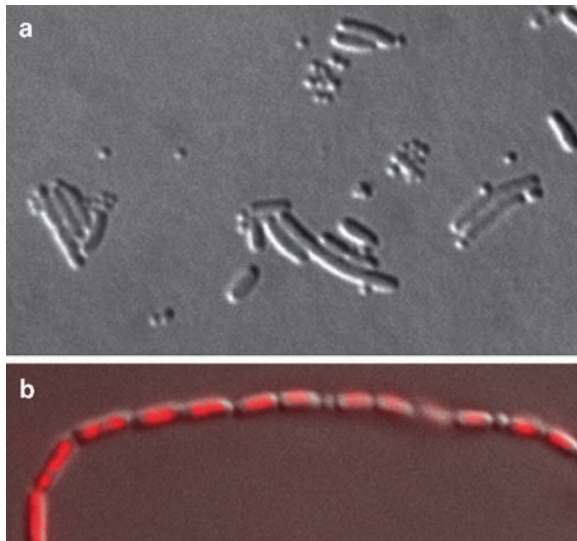


Fig. 2.5 DNA-free minicells caused by inactivating the Min system. *E. coli* cells lacking the Min proteins were imaged using DIC (**a**). Panel (**b**) shows a DIC/DAPI overlay of a cell chain of an *ftsK min* double mutant that contains DNA-free minicells within the chain. DAPI fluorescence is pseudocoloured red

divide again if the division block is removed). The striking observation is that nucleoids are usually evenly spaced throughout the filamentous cells, indicating that their spatial distribution is controlled (Fig. 2.6). Certain proteins, including FtsZ, but also others that normally localize to poles such as chemoreceptors and *Shigella* IcsA, localize between the nucleoids in nondividing cells, indicating that there are iterated markers for future cell poles at regular intervals in these filaments (Janakiraman and Goldberg 2004; Thiem et al. 2007). The forces of transertion (see below) may be another mechanism that provides spatial balance to centre the nucleoid.

Just as the nucleoid influences positioning of the cytokinetic apparatus, we know that the cytokinetic apparatus in turn can direct nucleoid positioning. In *E. coli*, one component of the cell division machinery is FtsK, a very large protein of 1,329 amino acids that colocalizes with the Z ring. However, only the first 200 amino acids are required for cell division (Yu et al. 1998a). The remainder of the protein harbours an ATPase-driven motor that transports the replication terminus portion of the chromosome, which is the last to be duplicated, to either side of the developing cell division septum (Bigot et al. 2007). This is a second checkpoint, after the SlmA/Noc checkpoint, to prevent nucleoid bisection by the division septum. In the absence of this checkpoint, nucleoids in a subset of cells are inappropriately positioned with respect to the septum, often resulting in frequent guillotining of the nucleoids and inhibition of septation (Fig. 2.7). Interestingly, despite the nucleoids in these division-defective cells being out of register with division septa, they still retain their normal separation and distribution, further supporting a dedicated

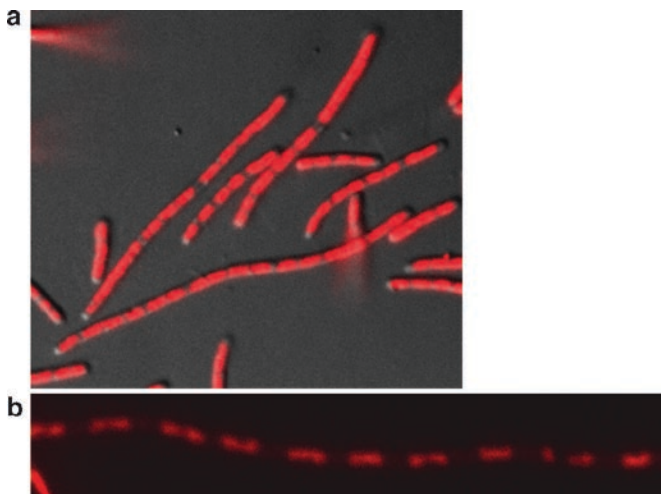


Fig. 2.6 Nucleoid positioning is independent of cytokinesis. Cytokinesis of *E. coli* cells was blocked by inactivation of FtsZ, and the nucleoids in the resulting nonseptated filamentous cells were visualized either after fixation and staining with DAPI (a) or by production of the *B. subtilis* Noc protein fused to cyan fluorescent protein (CFP) and microscopy of a live nondividing cell (b). Both images are overlays of DIC and DAPI fluorescence (pseudocoloured red). Chromosomal replication and partitioning functions appear to be mostly normal in these filamentous cells

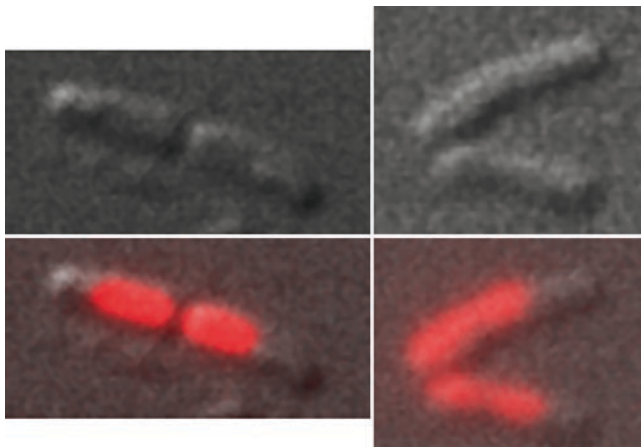


Fig. 2.7 Inappropriate positioning of nucleoids relative to the division septum in cells lacking the C terminus of FtsK leads to nucleoid guillotining. Typical cells lacking the C terminus of FtsK were imaged by DIC (*top*) or DAPI fluorescence (*bottom*). DAPI fluorescence is pseudocoloured red

nucleoid distribution system as discussed above. In *B. subtilis* undergoing sporulation, a homologue of FtsK, called SpoIIIE, helps to transport one of the daughter nucleoids through a pore in the sporulation septum into the prespore (Bath et al. 2000). In *E. coli*, FtsK has an additional function of stimulating the resolution of chromosome dimers that result from recombination events during replication (Aussel et al. 2002). This dimer resolution is also crucial to prevent guillotining of a trapped nucleoid containing an unpartitioned chromosome.

2.1.3 Factors that Shape the Nucleoid

The nucleoid of *E. coli* has a typical elongated shape during rapid growth. It occupies a large percentage of the cell volume but is sufficiently compact that the cell poles are nucleoid-free. Nucleoid duplication and partitioning results in a nucleoid-free zone at the cell centre, which allows the Z ring to form. Nucleoids in other bacteria generally have similar properties, although as mentioned above, a nucleoid-free zone at midcell is not always a prerequisite for Z ring formation. Nucleoid morphology is clearly regulated by a number of factors, because mutations and other perturbations to the cell can radically alter the degree of nucleoid compactness.

2.1.3.1 Effects of Transcription and Translation

Transcription and translation activity cause significant changes in nucleoid compaction. Inhibition of RNA polymerase activity by treatment with the drug rifampicin causes nucleoids to lose their relatively compact organization and fill virtually the entire cell volume (Dworsky and Schaechter 1973) (Fig. 2.8). This is not a

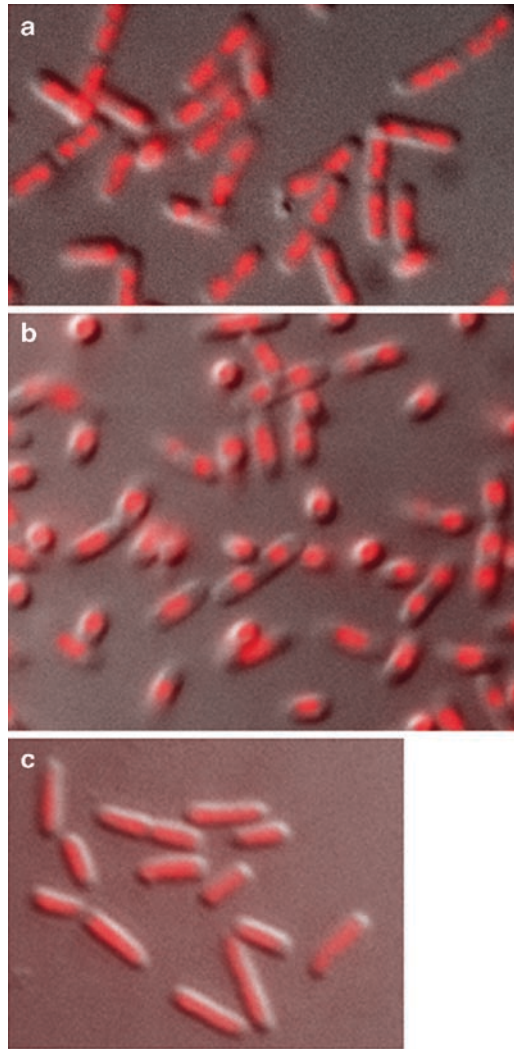


Fig. 2.8 Effects of inhibiting transcription or translation on nucleoid morphology. *E. coli* cells were either untreated (**a**), treated with chloramphenicol to block protein synthesis (**b**) or rifampicin to block transcription (**c**), then fixed and stained with DAPI (pseudocoloured red). Images shown are DIC/fluorescence overlays

nonspecific effect of rifampicin, because inhibition of the activity of σ^{70} by thermoinactivation of a conditional mutant allele has a similar effect, as does thermoinactivation of an RNA polymerase core subunit (Kruse et al. 2006; Sun and Margolin 2004). The requirement of transcription for nucleoid shape suggests that synthesis of rRNA, but also possibly other specific RNAs, are somehow involved in maintaining the integrity of the nucleoid mass. One possibility is that RNA polymerases themselves act as motors to continuously remodel bacterial chromatin and

keep the nucleoid from expanding outward. Another is that small RNAs themselves are involved in folding and compacting chromatin (Ohniwa et al. 2007). Other factors also help to compact the nucleoid (see below). Interestingly, rifampicin-induced decondensed nucleoids suppress the nucleoid occlusion effect, allowing FtsZ to assemble extensively in all areas of the cell (Sun and Margolin 2004). This suggests that nucleoid occlusion by Noc and SlmA is proportional to the local protein concentration, which in turn is mediated by the local density of DNA.

In contrast, inhibiting translation has the opposite effect on the nucleoid. Treatment of *E. coli* with the translational inhibitor chloramphenicol causes nucleoids to become significantly more compact compared to no drug treatment (Fig. 2.8). In fact, chloramphenicol can induce separated nucleoids to fuse (van Helvoort et al. 1996). These findings suggest that translation coupled with insertion of proteins into the membrane keeps the nucleoid engaged with the cytoplasmic membrane and thus in an expanded form (Woldringh 2002). Another possibility is that the more extended form of the nucleoid requires some other product of translation or the ribosomes themselves. However, it was shown recently that *E. coli* starved with serine hydroxamate, which induces production of the alarmone ppGpp and should also decrease translation activity, contained nucleoids that were decondensed compared to those of untreated cells (Ferullo and Lovett 2008). Similar decondensed nucleoids were observed when ppGpp levels were increased artificially by increasing levels of the RelA enzyme that synthesizes ppGpp. It is possible that a low threshold level of translation is sufficient to maintain nucleoid-membrane contacts needed for an expanded nucleoid. Interestingly, this same study showed that starved cells that lacked the ability to synthesize ppGpp and to mount a stringent response had highly condensed nucleoids, similar to those treated with chloramphenicol. Therefore, ppGpp seems to have a nucleoid-decondensing role, possibly by reducing rRNA transcription.

2.1.3.2 SMC Proteins and Nucleoid Condensation

SMC (Structural maintenance of chromosomes) proteins are conserved from bacteria to humans and function in large-scale organization of chromosomes (Hirano 2006). Eukaryotes contain four SMC variants, two of which control sister chromosome cohesion while the other two control chromosome condensation. Bacteria generally have a single SMC protein for chromosome condensation. An SMC monomer consists of a long coiled coil domain capped by a terminal globular ATP-binding domain. The functional form of SMC is a homodimer attached via the two globular domains, creating a flexible V-shaped molecule that is thought to act like a pincer to pack DNA when stimulated by other proteins. In the absence of SMC (or the MukBEF complex in *E. coli*, of which MukB is the SMC homologue), the nucleoids become more diffuse and mislocalized (Niki et al. 1992). Because cell division still occurs under these conditions, many nucleoids are not partitioned properly to progeny cells, resulting in a strikingly high percentage of cells lacking nucleoids (Fig. 2.9). These differ from minicells in that most of the nucleoid-free cells are as large as normal-sized cells.

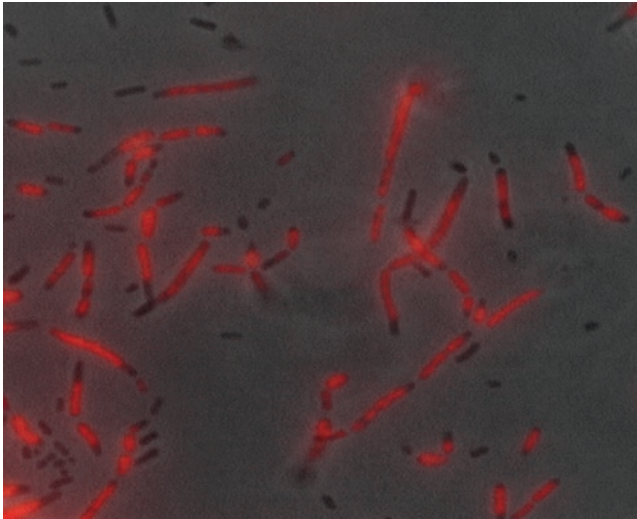


Fig. 2.9 High percentage of cells lacking nucleoids resulting from chromosome partitioning defects. *E. coli* cells with a deletion of the *mukB* gene were grown at 37°C for 2 h prior to fixation and staining with DAPI. Shown is an overlay of phase-contrast and DAPI fluorescence channels; DAPI fluorescence is pseudocoloured red. Note the many nucleoid-free cells lacking red fluorescence

If the transport functions of FtsK or SpoIIIE are blocked, then inactivation of SMC or MukB becomes lethal, presumably because both nucleoid decondensation coupled with lack of transport away from the septum results in guillotining of most nucleoids by the division septum (Britton and Grossman 1999; Yu et al. 1998b). In *E. coli*, the packing defect resulting from a lack of the MukBEF complex can be largely suppressed by inactivating Topoisomerase I (encoded by the *topA* gene), which increases negative chromosomal DNA supercoiling (Sawitzke and Austin 2000). This clearly demonstrates that supercoiling affects nucleoid compaction. However, the normal separation of the two arms of the chromosome in the two daughter cells remains defective in *mukB topA* double mutants (Danilova et al. 2007). This indicates that packing alone is not sufficient for proper nucleoid organization, although it is an important factor.

2.1.3.3 Other Factors that Shape the Nucleoid

Bacterial chromatin is not nearly as packed with protein compared with eukaryotic chromatin, which is wound around histones. However, bacteria have a number of nucleoid-associated proteins (NAPs) that influence nucleoid organization. For example, altering one of these proteins in *E. coli*, HU, dramatically increases nucleoid compaction (Kar et al. 2005). Similar effects are observed

upon overexpression of H-NS (Spurio et al. 1992). Other NAPs, which will be described below and in more detail in later chapters, also have effects on nucleoid condensation.

Protection of the chromosomal DNA seems to be correlated with nucleoid compaction. For example, the Dps protein helps to condense the *E. coli* nucleoid into a crystalline assembly during stationary phase (Frenkiel-Krispin et al. 2004; Ohniwa et al. 2006). In *Staphylococcus aureus*, a Dps-like protein condenses the nucleoid not in stationary phase but specifically during oxidative stress (Morikawa et al. 2006). Nucleoids of bacteria that are highly resistant to radiation-induced DNA damage, such as *Deinococcus* species, are characterized by their highly condensed nucleoids (Zimmerman and Battista 2005). The mechanism of radioresistance and its correlation to a compact nucleoid is not yet resolved.

Other NAPs seem to have housekeeping functions in organizing the chromosome. CspE, originally found as a multicopy suppressor of the chromatin-disrupting agent camphor in *E. coli*, also suppresses decondensation defects caused by inhibition of MukB (Yamanaka et al. 1994). CspE may help to condense the nucleoid by binding different DNA segments followed by formation of CspE homodimers, potentially via a mechanism similar to SMC proteins (Johnston et al. 2006). Other proteins, including SeqA, which is involved in organizing the nucleoid during chromosome replication and segregation, also help to sculpt the nucleoid (Weitao et al. 1999). It should be emphasized that non-protein factors are also important for nucleoid compaction. They include macromolecular crowding effects and polyamines (Sarkar et al. 2009; Zimmerman 2006).

Nucleoid morphologies sometimes change significantly during bacterial differentiation. During its life cycle, *Chlamydia* species differentiate from elementary bodies (EBs) to reticulate bodies (RBs). EBs are the very small, infectious but otherwise inert form of this obligate intracellular pathogen. They carry a highly condensed form of chromatin (Costerton et al. 1976). Once inside the host cell, EBs differentiate into RBs, which are metabolically active, larger than EBs, and divide by binary fission. During this transition, the nucleoids become decondensed. Finally, the multiple RBs differentiate back into EBs, and the nucleoids recondense. The high degree of nucleoid compaction in EBs depends on a chlamydial NAP, called Hc1 because of its similarity to eukaryotic histone H1. In fact, when it is synthesized in *E. coli*, Hc1 is sufficient to compact the *E. coli* nucleoid to the point of lethality (Barry et al. 1992). Similar NAP-mediated changes in nucleoid compactness occur in other bacteria that differentiate, including *Coxiella burnetii* (Heinzen et al. 1996).

Another example of large changes in nucleoid morphology during differentiation is the process of endospore formation and germination in *Bacillus* species. During the first stages of endospore formation after starvation of vegetative *B. subtilis*, the chromosomes undergo a final round of duplication but transiently form a narrow extended structure called the axial filament (Bylund et al. 1993). This filament, attached at both poles by RacA (Ben-Yehuda et al. 2003), helps to pull the spore-bound nucleoid poleward into the developing prespore compartment.

The nucleoid of the spore will become highly compact and complexed with small acid-soluble proteins (SASPs) for protection of the DNA (Lee et al. 2008). Germination of the spore results in further changes in nucleoid morphology. The nucleoids of spores from *Bacillus megaterium* form tight ring structures immediately after germination, when the SASPs are still DNA-bound (Ragkousi et al. 2000). Within 15 min of germination, the nucleoid condenses further, rendering the lumen of the previous ring structure invisible. This step correlates with degradation of the SASPs. Finally, after several hours of outgrowth from the spore coat, the cells elongate and the nucleoids attain their usual lobed appearance.

2.1.4 Special Cases

Most of this chapter has pertained to nucleoids in bacteria that divide by binary fission, as they are the best characterized. However, not surprisingly, there is great diversity in overall nucleoid organization among diverse bacteria. For example, *Gemmata obscuriglobus*, a member of the highly diverse Planctomycetes phylum, divides by budding, as do a variety of other bacteria. But what distinguishes this and related species is the compartmentalization of their nucleoids. Indeed, their highly condensed nucleoids are bounded by an intracytoplasmic membrane (Lee et al. 2009). When a daughter bud is initiated off the mother cell, it receives a “naked” nucleoid, and then surrounds it by a membrane as the bud matures.

Epulopiscium, on the other hand, is most notable for being one of the largest bacteria (>200 μm long) and for giving birth to live progeny from within. In addition, as might be expected from the large cytoplasmic mass, each *Epulopiscium* cell harbours hundreds of separate nucleoids around its periphery, while the rest of the cell remains DNA-free (Mendell et al. 2008). The result is an extreme case of polyploidy, more reminiscent of eukaryotes than prokaryotes.

2.1.5 Outlook

The bacterial nucleoid has been observed since the early twentieth century, and recent advances in bacterial cytology have elucidated much about the internal organization of the nucleoid as well as determinants of its shape under various conditions. It is still not clear how nucleoids retain their characteristic shapes and positions within the cell, how NAPs regulate these properties, or how partitioning to daughter cells occurs. As most studies of the nucleoid have been done in only a few model systems, it is likely that we will learn more from investigating diverse species.

References

- Adler HI, Fisher W, Cohen A, Hardigree A (1967) Miniature *E. coli* cells deficient in DNA. *Proc Natl Acad Sci USA* 57:321–326
- Aussel L, Barre FX, Aroyo M, Stasiak A, Stasiak AZ, Sherratt D (2002) FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* 108:195–205
- Barry CE 3rd, Hayes SF, Hackstadt T (1992) Nucleoid condensation in *Escherichia coli* that express a chlamydial histone homolog. *Science* 256:377–379
- Bath J, Wu LJ, Errington J, Wang JC (2000) Role of *Bacillus subtilis* SpoIIIE in DNA transport across the mother cell-prespore division septum. *Science* 290:995–997
- Bendich AJ, Drlica K (2000) Prokaryotic and eukaryotic chromosomes: what's the difference? *Bioessays* 22:481–486
- Ben-Yehuda S, Rudner DZ, Losick R (2003) RacA, a bacterial protein that anchors chromosomes to the cell poles. *Science* 299:532–536
- Bernhardt TG, de Boer PA (2005) SlmA, a nucleoid-associated, FtsZ-binding protein required for blocking septal ring assembly over chromosomes in *E. coli*. *Mol Cell* 18:555–564
- Bi E, Lutkenhaus J (1991) FtsZ ring structure associated with division in *Escherichia coli*. *Nature* 354:161–164
- Bigot S, Sivanathan V, Possoz C, Barre FX, Cornet F (2007) FtsK, a literate chromosome segregation machine. *Mol Microbiol* 64:1434–1441
- Bohrmann B, Villiger W, Johansen R, Kellenberger E (1991) Coralline shape of the bacterial nucleoid after cryofixation. *J Bacteriol* 173:3149–3158
- Britton RA, Grossman AD (1999) Synthetic lethal phenotypes caused by mutations affecting chromosome partitioning in *Bacillus subtilis*. *J Bacteriol* 181:5860–5864
- Bylund JE, Haines MA, Piggot PJ, Higgins ML (1993) Axial filament formation in *Bacillus subtilis*: induction of nucleoids of increasing length after addition of chloramphenicol to exponential-phase cultures approaching stationary phase. *J Bacteriol* 175:1886–1890
- Costerton JW, Poffenroth L, Wilt JC, Kordova N (1976) Ultrastructural studies of the nucleoids of the pleomorphic forms of *Chlamydia psittaci* 6BC: a comparison with bacteria. *Can J Microbiol* 22:16–28
- Danilova O, Reyes-Lamothe R, Pinskaya M, Sherratt D, Possoz C (2007) MukB colocalizes with the *oriC* region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol Microbiol* 65:1485–1492
- Dworsky P, Schaechter M (1973) Effect of rifampin on the structure and membrane attachment of the nucleoid of *Escherichia coli*. *J Bacteriol* 116:1364–1374
- Eltsov M, Dubochet J (2005) Fine structure of the *Deinococcus radiodurans* nucleoid revealed by cryoelectron microscopy of vitreous sections. *J Bacteriol* 187:8047–8054
- Eltsov M, Zuber B (2006) Transmission electron microscopy of the bacterial nucleoid. *J Struct Biol* 156:246–254
- Ferullo DJ, Lovett ST (2008) The stringent response and cell cycle arrest in *Escherichia coli*. *PLoS Genet* 4:e1000300
- Frenkiel-Krispin D, Ben-Avraham I, Englander J, Shimoni E, Wolf SG, Minsky A (2004) Nucleoid restructuring in stationary-state bacteria. *Mol Microbiol* 51:395–405
- Harry EJ, Rodwell J, Wake RG (1999) Co-ordinating DNA replication with cell division in bacteria: a link between the early stages of a round of replication and mid-cell Z ring assembly. *Mol Microbiol* 33:33–40
- Heinzen RA, Howe D, Mallavia LP, Rockey DD, Hackstadt T (1996) Developmentally regulated synthesis of an unusually small, basic peptide by *Coxiella burnetii*. *Mol Microbiol* 22:9–19
- Hinnebusch BJ, Bendich AJ (1997) The bacterial nucleoid visualized by fluorescence microscopy of cells lysed within agarose: comparison of *Escherichia coli* and spirochetes of the genus *Borrelia*. *J Bacteriol* 179:2228–2237

- Hirano T (2006) At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol* 7:311–322
- Janakiraman A, Goldberg MB (2004) Evidence for polar positional information in *E. coli* independent of cell division and nucleoid occlusion. *Proc Natl Acad Sci USA* 101:835–840
- Johnston D, Tavano C, Wickner S, Trun N (2006) Specificity of DNA binding and dimerization by CspE from *Escherichia coli*. *J Biol Chem* 281:40208–40215
- Kar S, Edgar R, Adhya S (2005) Nucleoid remodeling by an altered HU protein: reorganization of the transcription program. *Proc Natl Acad Sci USA* 102:16397–16402
- Kruse T, Blagoev B, Lobner-Olesen A, Wachi M, Sasaki K, Iwai N, Mann M, Gerdes K (2006) Actin homolog MreB and RNA polymerase interact and are both required for chromosome segregation in *Escherichia coli*. *Genes Dev* 20:113–124
- Lee KS, Bumbaca D, Kosman J, Setlow P, Jedrzejewski MJ (2008) Structure of a protein-DNA complex essential for DNA protection in spores of *Bacillus* species. *Proc Natl Acad Sci USA* 105(8):2806–2811
- Lee KC, Webb RI, Fuerst JA (2009) The cell cycle of the planctomycete *Gemmata obscuriglobus* with respect to cell compartmentalization. *BMC Cell Biol* 10:4
- Mason DJ, Powelson DM (1956) Nuclear division as observed in live bacteria by a new technique. *J Bacteriol* 71:474–479
- Mendell JE, Clements KD, Choat JH, Angert ER (2008) Extreme polyploidy in a large bacterium. *Proc Natl Acad Sci USA* 105:6730–6734
- Morikawa K, Ohniwa RL, Kim J, Maruyama A, Ohta T, Takeyasu K (2006) Bacterial nucleoid dynamics: oxidative stress response in *Staphylococcus aureus*. *Genes Cells* 11:409–423
- Morlot C, Zapun A, Dideberg O, Vernet T (2003) Growth and division of *Streptococcus pneumoniae*: localization of the high molecular weight penicillin-binding proteins during the cell cycle. *Mol Microbiol* 50:845–855
- Niki H, Hiraga S (1998) Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev* 12:1036–1045
- Niki H, Imamura R, Kitaoka M, Yamanaka K, Ogura T, Hiraga S (1992) *E. coli* MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. *EMBO J* 11:5101–5109
- Niki H, Yamaichi Y, Hiraga S (2000) Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14:212–223
- Ohniwa RL, Morikawa K, Kim J, Ohta T, Ishihama A, Wada C, Takeyasu K (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. *EMBO J* 25:5591–5602
- Ohniwa RL, Morikawa K, Takeshita SL, Kim J, Ohta T, Wada C, Takeyasu K (2007) Transcription-coupled nucleoid architecture in bacteria. *Genes Cells* 12:1141–1152
- Piekarski G (1937) Zytologische Untersuchungen an *Paratyphus*-und *Coli* bacterien. *Arch Mikrobiol* 8:428–429
- Ragkousi K, Cowan AE, Ross MA, Setlow P (2000) Analysis of nucleoid morphology during germination and outgrowth of spores of *Bacillus* species. *J Bacteriol* 182:5556–5562
- Ramos A, Letek M, Campelo AB, Vaquera J, Mateos LM, Gil JA (2005) Altered morphology produced by *ftsZ* expression in *Corynebacterium glutamicum* ATCC 13869. *Microbiology* 151:2563–2572
- Robinow C, Kellenberger E (1994) The bacterial nucleoid revisited. *Microbiol Rev* 58:211–232
- Rothfield L, Taghbalout A, Shih YL (2005) Spatial control of bacterial division-site placement. *Nat Rev Microbiol* 3:959–968
- Sarkar T, Petrov AS, Vitko JR, Santai CT, Harvey SC, Mukerji I, Hud NV (2009) Integration host factor (IHF) dictates the structure of polyamine-DNA condensates: implications for the role of IHF in the compaction of bacterial chromatin. *Biochemistry* 48:667–675
- Sawitzke JA, Austin S (2000) Suppression of chromosome segregation defects of *Escherichia coli* *muk* mutants by mutations in topoisomerase I. *Proc Natl Acad Sci USA* 97:1671–1676
- Spurio R, Durrenberger M, Falconi M, La Teana A., Pon CL, Gualerzi CO (1992) Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. *Mol Gen Genet* 231:201–211

- Stempen H (1950) Demonstration of the chromatinic bodies of *Escherichia coli* and *Proteus vulgaris* with the aid of the phase contrast microscope. J Bacteriol 60:81–87
- Sun Q, Margolin W (2001) Influence of the nucleoid on placement of FtsZ and MinE rings in *Escherichia coli*. J Bacteriol 183:1413–1422
- Sun Q, Margolin W (2004) Effects of perturbing nucleoid structure on nucleoid occlusion-mediated toporegulation of FtsZ ring assembly. J Bacteriol 186:3951–3959
- Teleman AA, Graumann PL, Lin DCH, Grossman AD, Losick R (1998) Chromosome arrangement within a bacterium. Curr Biol 8:1102–1109
- Thanbichler M, Shapiro L (2006) MipZ, a spatial regulator coordinating chromosome segregation with cell division in *Caulobacter*. Cell 126:147–162
- Thiem S, Kentner D, Sourjik V (2007) Positioning of chemosensory clusters in *E. coli* and its relation to cell division. EMBO J 26:1615–1623
- van Helvoort JM, Kool J, Woldringh CL (1996) Chloramphenicol causes fusion of separated nucleoids in *Escherichia coli* K-12 cells and filaments. J Bacteriol 178:4289–4293
- Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. Proc Natl Acad Sci USA 101:9257–9262
- Webb CD, Teleman A, Gordon S, Straight A, Belmont A, Lin DC, Grossman AD, Wright A, Losick R (1997) Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. Cell 88:667–674
- Weitao T, Nordström K, Dasgupta S (1999) Mutual suppression of *mukB* and *seqA* phenotypes might arise from their opposing influences on the *Escherichia coli* nucleoid structure. Mol Microbiol 34:157–168
- Woldringh CL (1976) Morphological analysis of nuclear separation and cell division during the life cycle of *Escherichia coli*. J Bacteriol 125:248–257
- Woldringh CL (2002) The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. Mol Microbiol 45:17–29
- Woldringh C, Mulder E, Valkenburg J, Wientjes F, Zaritsky A, Nanninga N (1990) Role of the nucleoid in the toporegulation of division. Res Microbiol 141:39–49
- Wu LJ, Errington J (2004) Coordination of cell division and chromosome segregation by a nucleoid occlusion protein in *Bacillus subtilis*. Cell 117:915–925
- Yamaichi Y, Niki H (2004) *migS*, a *cis*-acting site that affects bipolar positioning of *oriC* on the *Escherichia coli* chromosome. EMBO J 23:221–233
- Yamanaka K, Mitani T, Ogura T, Niki H, Hiraga S (1994) Cloning, sequencing, and characterization of multicopy suppressors of a *mukB* mutation in *Escherichia coli*. Mol Microbiol 13:301–312
- Yu X-C, Margolin W (1999) FtsZ ring clusters in *min* and partition mutants: role of both the Min system and the nucleoid in regulating FtsZ ring localization. Mol Microbiol 32:315–326
- Yu X-C, Tran AH, Sun Q, Margolin W (1998a) Localization of cell division protein FtsK to the *Escherichia coli* septum and identification of a potential N-terminal targeting domain. J Bacteriol 180:1296–1304
- Yu XC, Weihe EK, Margolin W (1998b) Role of the C terminus of FtsK in *Escherichia coli* chromosome segregation. J Bacteriol 180:6424–6428
- Zimmerman SB (2003) Underlying regularity in the shapes of nucleoids of *Escherichia coli*: implications for nucleoid organization and partition. J Struct Biol 142:256–265
- Zimmerman SB (2006) Cooperative transitions of isolated *Escherichia coli* nucleoids: implications for the nucleoid as a cellular phase. J Struct Biol 153:160–175
- Zimmerman JM, Battista JR (2005) A ring-like nucleoid is not necessary for radioresistance in the *Deinococcaceae*. BMC Microbiol 5:17

Chapter 3

The Chromosome Segregation Machinery in Bacteria

Peter L. Graumann

Abstract Two different mechanisms for chromosome segregation can be envisioned in bacteria based on current knowledge. In some bacteria that have a single chromosome, the separation of sister chromosomes is achieved by an active machinery that can move individual regions on the chromosomes over a distance of several microns within a few minutes. This process is independent of cell elongation. Several key factors have been identified for this active machinery and will be discussed. However, the driving “motor” has not yet been discovered, although suspicious candidates like filament-forming proteins and RNA polymerase are under intensive investigation. Alternatively, chromosomes can be randomly separated if several copies are present within the cell, which seems to occur in the bacterial phylum of Cyanobacteria, and possibly in many other species.

Keywords Chromosome segregation • bacterial cell cycle • partitioning • topoisomerase • structural maintenance of chromosomes

Introduction

No matter how plausible at the time, the old Jacob/Brenner/Cazin model for cell growth-driven chromosome segregation (Jacob et al. 1963) is no longer accepted. Some text books still state that parts of the chromosome are at least transiently attached to the cell membrane. However, this concept has only been proven in the special case of attachment of chromosome origins to the cell poles in sporulating *Bacillus subtilis* cells (Ben-Yehuda and Losick 2002), but not for any other system in any clearly defined manner. Instead, highly dynamic movements of individual regions on the chromosomes either from the cell centre towards opposite cell poles or from one pole to the other pole have been visualized in many bacteria, which

P.L. Graumann (✉)

Microbiology, Faculty for Biology, University of Freiburg, Schänzle Straße 1, 79104, Freiburg, Germany

e-mail: peter.graumann@biologie.uni-freiburg.de

cannot be explained by slow extension of the cell wall and passive co-migration of the chromosome. The movement of chromosome regions can be visualized through the integration of an array of lactose operators (about 250) into a region on the chromosome, which is subsequently decorated by lactose repressor-GFP fusions (LacI-GFP) (Webb et al. 1997). Alternatively, ParB-type protein-GFP fusions that bind to and spread away from ParB binding sites have been employed. LacI-GFP decorated origin regions in *B. subtilis* and in *Escherichia coli* typically move from the cell centre, where they are duplicated, towards opposite cell poles within 2–4 min (over a distance of about 1.5 μm), with peak speeds of about 0.3 $\mu\text{m}/\text{min}$ (Gordon et al. 1997; Webb et al. 1998). This movement was also monitored for the terminus region on the chromosome, as well as for sites in between origin and terminus regions, and occurs even during inhibition of cell wall synthesis (and thus cell extension) (Webb et al. 1998). In *Caulobacter crescentus*, the origin region localizes to one of the cell poles, but after initiation of replication close to this pole, one of the duplicated origin regions rapidly moves across the cell to the other cell pole (Viollier et al. 2004) (Fig. 3.1b). Additionally, chromosome segregation follows a highly ordered pattern, with regions that are duplicated later being segregated later than earlier ones, or in other words, segregation of chromosome regions follows the temporal pattern in which they are duplicated (Teleman et al. 1998) (Fig. 3.2). In *E. coli*, a time of cohesion of chromosome regions appears to exist, because chromosome origins remain close to each other for an extended time,

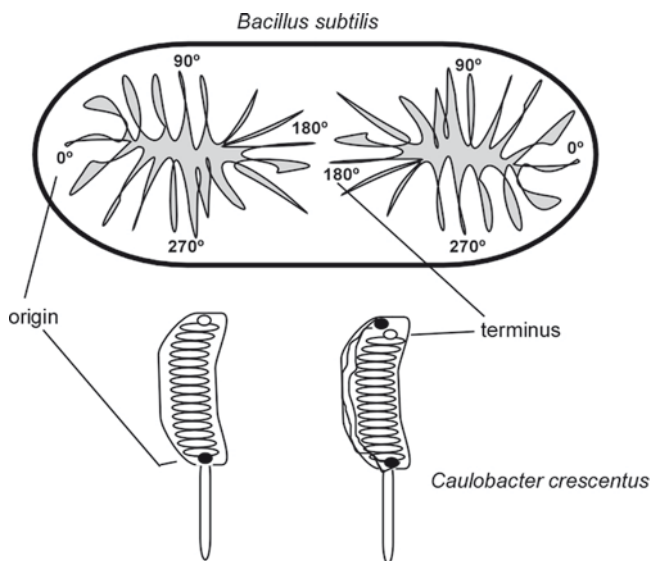


Fig. 3.1 Chromosome segregation and arrangement patterns in different bacteria. (a) Preferred arrangement of the chromosome in *B. subtilis*. Chromosome origins are termed 0°, and terminus regions 180°. The mechanism that sets up the arrangement during the cell cycle is depicted in Fig. 3.2. (b) Initial events in chromosome segregation in *Caulobacter crescentus*. Following initiation of replication, one duplicated origin moves across the cell to the other cell pole. The chromosome is drawn as many connected DNA loops

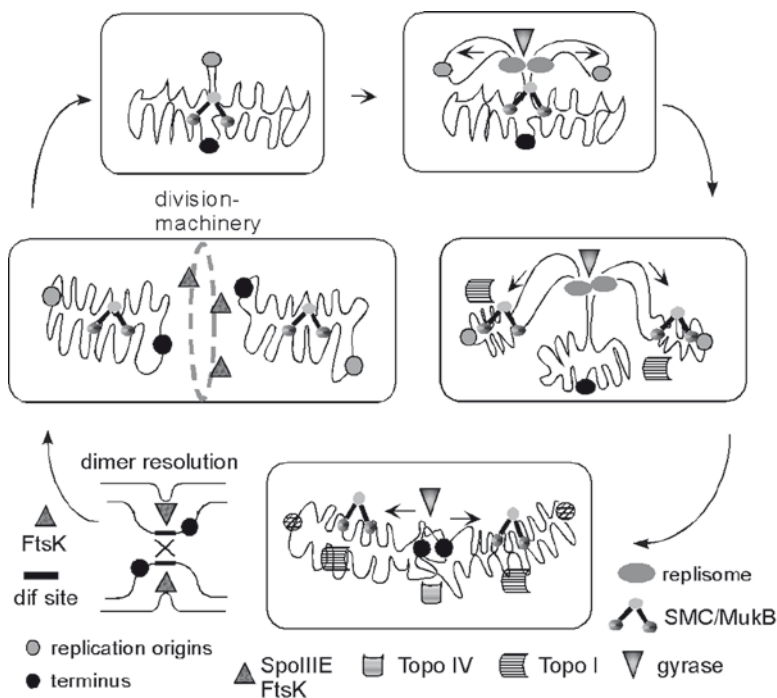


Fig. 3.2 Cell cycle regulated events during chromosome segregation in *B. subtilis*. Arrows indicate the direction of segregation of duplicated chromosome regions

together with several regions quite far away from the origins (Bates and Kleckner 2005; Sunako et al. 2001). After this period of apparent cohesion (which could also be based on active segregation following initiation of replication much later than in e.g. *B. subtilis* cells), origins and other already duplicated regions separate rather synchronously towards opposite cell poles. Thereafter segregation of newly replicated regions occurs much sooner after their duplication compared with origin regions. In toto, these findings show that an active intracellular segregation machinery exists in bacteria that drives or help drive and coordinates this vital cell cycle process. Several interesting studies have pointed out that the DNA itself has physiological factors that may help driving apart the sister chromosomes. For example, the proposed pulling of DNA to the extending cell membrane and physiological factors such as repulsion of the negative charges in DNA and molecular crowding within the cells may facilitate the segregation of duplicated chromosome regions (Woldringh and Nanninga 2006), which is discussed in detail in Chapter 5. However, these accessory forces cannot by themselves explain the highly ordered pattern of chromosome segregation found in a variety of bacteria. The known components of the segregation machinery will be discussed in the next sections.

On the other hand, many bacteria contain several copies of the chromosome, such that by analogy to the segregation of high copy number plasmids, chromosomes could be separated at random, and the high number of chromosomes would

ensure that the daughter cells obtain a reasonable number of chromosomes by default. This non-stringently regulated mechanism appears to occur in some bacteria, and will be discussed in the last part of this chapter.

3.1 Active Chromosome Segregation

The bacterial counterpart of the eukaryotic mitotic segregation machinery is under heavy investigation. Bacteria do not appear to contain a tubulin-based spindle apparatus (rather, the bacterial tubulin ortholog FtsZ drives cell division; Margolin 2005) or kinesin/dynein motor proteins as found in eukarya. However, some segregation proteins are conserved between bacteria and eukarya, amongst these topoisomerases and SMC proteins, which will be discussed below. Two key findings on chromosome segregation were made about 10 years ago. Firstly, it was shown for several organisms that the chromosome has a preferred ordered arrangement within the cell. This arrangement is apparently different between different species, but in general, the bacterial chromosome is arranged roughly according to its physical structure. In *B. subtilis*, origin regions on the chromosome are separated early in the cell cycle, soon after their duplication, and remain positioned close to the cell poles for the rest of the cell cycle (Webb et al. 1997). The terminus region is located at the cell centre and the duplicated regions are moved apart just before cell division occurs. Positions between origin and terminus are also located between the poles and the cell centre (Fig. 3.1a) (Teleman et al. 1998). Similarly, the chromosome has a preferred and ordered arrangement in *E. coli* cells (Niki et al. 2000). In *C. crescentus*, the origin is located at the stalked cell pole, and the terminus at the other cell pole. Again, all other regions on the chromosome are positioned at intermediate positions according to a defined order (Viollier et al. 2004) (Fig. 3.1b). After initiation of replication, one origin moves across the entire cell and becomes anchored at the opposite cell pole. All consequently duplicated chromosome regions become translocated into the other cell half, such that inversely oriented but well organized chromosomes are set up in each daughter cell (Fig. 3.1b). Thus, overall, the order of genes on the chromosomes roughly reflects where in the cell the genes are positioned, of course with some degree of flexibility.

Secondly, it was found that the replication machinery – i.e. both replication forks – is located near the cell centre in *B. subtilis* and in *E. coli*, or at the cell pole in *C. crescentus*, at the beginning of the cell cycle (and after initiation of replication). In *B. subtilis*, both forks remain close to the centre, until later in the cell cycle, the forks move to the quarter sites, and remain there until termination of replication takes place (Lemon and Grossman 1998). In *E. coli*, the two forks slowly move towards opposite cell poles (Reyes-Lamothe et al. 2008), while in *C. crescentus*, the replisome moves slowly from one pole towards the cell centre, where it terminates replication (Jensen et al. 2005). None of the replication machineries appear to be anchored at a specific site. For example, the *B. subtilis* replication machinery can be seen as two foci, most likely the two forks, which move away from each other and reunite in a matter of seconds to minutes (Peters et al. 2007). Although both

forks are very close to each other throughout the largest part of the cell cycle (i.e. they are seen as a single focus), they operate independently from each other, and can move very rapidly around the cell centre, or towards the cell poles. However, overall, it is clear that the replication forks do not run over the chromosome, but that the chromosome moves through the replication machinery, and duplicated regions are moved apart with little or even no delay, except for *E. coli*, where the first half of the chromosome is segregated with some delay relative to initiation of replication (Bates and Kleckner 2005). Thus, replication and segregation operate in an ordered fashion, and replicated regions on the chromosome are moved towards opposite cell poles in a coordinated manner during ongoing replication. This organization of replication greatly facilitates chromosome segregation, because only individual regions or domains of the chromosome need to be physically moved apart in a sequential and ongoing manner, and not the whole chromosome at a single moment in the cell cycle. After movement, segregated chromosome regions can be organized such that they rapidly obtain a relatively defined position within the cell, keeping chromosome arrangement in check, except for the short moments of movement of the regions through the replication machinery and their movement away from the replisome.

Under rapid growth conditions, many bacterial cells contain multiple origins and replication forks, such that less time is needed to get through the cell cycle. Cells divide with two copies of the chromosome being almost fully replicated, and several replication forks already being active, such that less time is required for a cell cycle – which can be as fast as 20 min doubling time – than is needed for the replication of the entire chromosome (about 40–50 min) (Fossum et al. 2007). Under these conditions, the arrangement of replication forks and chromosome origins (and the rest of the chromosome) is simply multiplied, i.e. multiple regularly arranged replication machineries spill out the regularly spaced chromosome origins and the following chromosome regions. So the segregation scheme in Fig. 3.2 would involve additional replication forks operating between segregated origin regions, but overall, ordered replication/segregation would be maintained.

How could the replication forks be kept close to the cell centre? In several bacteria, the actin-like MreB protein has been implicated in chromosome segregation. Mutations in *mreB* lead to the loss of proper cell growth and thus cell shape, and to the loss of cell polarity in many rod shaped or curved bacteria (Doi et al. 1988; Jones et al. 2001). Depletion of MreB or mutation of its ATPase activity leads to a strong defect in chromosome segregation (Defeu Soufo and Graumann 2003; Kruse et al. 2003). Even if MreB is not directly involved in chromosome segregation, its lack clearly disorganizes segregation, for example, origin regions frequently move to the same cell pole in *B. subtilis*. Interestingly, the *B. subtilis* replication machinery loses its central position during the depletion of MreB, although at this point, the cells still continue to grow as rods, and only later lose their cell shape (Defeu Soufo and Graumann 2005). It has been speculated that MreB sets up or facilitates the movement of duplicated chromosome regions towards opposite cell poles, and that the simultaneous pushing or pulling of duplicated chromosome segments simply keeps the replication machinery around the cell centre. In other words, there may be a tug of war for newly duplicated regions on the chromosome towards

opposite poles, and the net result is that the replisome remains positioned. In the absence of proper segregation, replication forks start to move along the chromosome and end up at random places within the cells.

3.1.1 *MreB*

The implication of MreB in chromosome segregation is still highly controversial. It has been shown that MreB sets up a cytoskeleton-like structure in many bacteria, because it forms filamentous structures that run in a helical pattern underneath the cell membrane (Jones et al. 2001). These filamentous structures may be polarized, based on the idea that MreB has a structure that is highly similar to actin, and forms actin-like filaments in vitro, which have a plus and a minus end. Polarized MreB filaments could be used to orient components of an active chromosome segregation machinery, e.g. unknown motor proteins, yet to be identified. A motor-like mechanism has been proposed based on the highly dynamic localization of MreB within the cells (Defeu Soufo and Graumann 2004). Extension of MreB polymers from the cell centre towards the cell poles (or from one pole to the other in *C. crescentus*) may push newly replicated DNA through the cells. In support of this idea, blocking of MreB polymerisation blocks segregation of origins in *C. crescentus*, and MreB interacts with origin DNA as shown by ChIP experiments (Gitai et al. 2005). Another very intriguing mechanism has been proposed based on the observation that MreB interacts with RNA polymerase (RNAP) (Kruse et al. 2006). This enzyme can exert a high force onto the DNA that is used as a template for transcription. If one imagines RNAP being kept in place, DNA is moved through the enzyme with a high force that exceeds the force exerted by e.g. kinesin. If MreB indeed anchors many RNAP molecules underneath the membrane, and orients them, then their combined action could push DNA rapidly and over a long distance. Intriguingly, the vast majority of genes is oriented away from the origin regions, such that transcription of these genes (which are usually also the most highly transcribed genes on the chromosome) through stationary RNAP molecules may move DNA towards the cell poles. Indeed, when transcription is blocked, even duplicated origin regions fail to get separated to the poles, showing that active transcription is required for chromosome segregation (Dworkin and Losick 2002). This model is highly attractive because it involves proteins that are present for an essential task, and does not involve any additional further specially evolved factors.

3.1.2 *ParA and ParB*

ParA type proteins belong to a family of so-called P-loop ATPases (which include MinD proteins) and are implicated in plasmid partitioning. Low copy number plasmids contain two known types of active partitioning systems: ParA/ParB system

(type I), and the ParR/ParM system (which is also called type II segregation system, where filaments formed by the actin-like ParM protein push plasmids towards opposite cell poles) (Chapter 4). The plasmid-encoded ParAB systems consist of a defined partitioning site (*par*), located close to the *parAB* bicistronic operon. ParB-type proteins bind to the *par* site (or in some cases multiple sites) situated close to the *par* operon, and interact with their ParA counterparts (Gerdes et al. 2004). ParA type proteins form helical filaments that rapidly oscillate along the nucleoids (Ebersbach and Gerdes 2004). In vitro, ParA forms ATP-dependent filaments and bundles of filaments, in which filaments appear to twist around each other (Barilla et al. 2005; Ebersbach et al. 2006; Kuempel et al. 1991). Consistent with the formation of ordered filamentous structures, ParA ATPase activity is cooperative. Thus, formation of dynamic filaments appears to be a conserved function of P-loop ATPases. The interaction between ParA and ParB suggests that ParA dynamics distribute plasmids along the nucleoids and thus ensure that sufficient copies are positioned within each cell half before cell division occurs. It is also interesting to note that a helical pattern is a recurring scheme for the localization of many bacterial proteins forming dynamic filaments.

Interestingly, a ParA protein has been shown to be involved in chromosome segregation in *Vibrio cholerae*, which contains a large (chrI) and a smaller (chrII) chromosome. ChrI is segregated similar to that of *C. crescentus*: the origin is located at one cell pole, and the newly duplicated origin is translocated across the cell to the other pole. In contrast, the origin of chrII is located near mid cell and duplicated origins move towards the quarter sites (Fogel and Waldor 2005), as in *B. subtilis*. Both chromosomes encode ParAB systems. While MreB has been implicated in the segregation of both chromosomes (Srivastava et al. 2007), segregation of chrI becomes abnormal in the absence of ParAI, in that origins are no longer tethered to the cell poles, but now segregate in a manner resembling that of origins of chrII (Fogel and Waldor 2006). ParAI forms a band that appears to consist of helical filaments from one cell pole to the other that contains the duplicated origins. ParAI also interacts with ParBI, which binds to sites close to the origin of chrI. Concomitant with the shrinkage of the ParAI band, one origin (through its interaction with ParBI) follows the retracting ParAI filaments towards the opposite cell pole, suggesting that ParAI is involved in a mechanism that pulls origin regions across the cell. However, it appears that the mechanism that partitions chrII can take over segregation of chrI (i.e. possibly MreB-based), because loss of ParAI does not lead to any discernable growth defect.

3.1.3 Topoisomerases

Two types of topoisomerases exist in bacteria. Type I enzymes generate a single strand nick in the DNA and move the non-broken strand through this nick – this changes the linking number by one. Topoisomerase I (Topo I) and Topo III are type

I enzymes, and Topo I is known to change the linking number towards positive (or rather less negative) values (Wang 2002; Zechiedrich et al. 2000), opposing the activity of Topo II (called DNA gyrase or simply gyrase). Like Topo IV, gyrase belongs to type II enzymes that generate double stranded gaps, through which another section of the double-stranded DNA is moved, changing the linking number by a value of 2 (Stone et al. 2003). Gyrase uses ATP to introduce negative supercoils into DNA, which opens up DNA strands to facilitate replication and transcription of DNA. In mesophilic bacteria (those growing between about 8°C and 50°C), gyrase sets up overall negative superhelical density in the chromosome (about 6% underwinding of the DNA), and the counteracting activity of Topo I ensures that this value is not increased. Thus, topoisomerases are essential for the maintenance of superhelical density and thus of vital cellular processes concerning the chromosome, including the compaction of DNA. In bacteria such as *B. subtilis* and *E. coli*, the chromosome forms a visible structure called the nucleoid, i.e. it is compacted in the centre of the cell, while the cell poles are devoid of DNA (Fig. 3.3b). In addition to its conserved arrangement, the nucleoid is organized into topological domains, that is entities, whose superhelical density is independent from each other (Postow et al. 2004; Staczek and Higgins 1998) (Chapter 7). About 400 of such domains exist, and as these domains are topologically isolated for example, a double stranded break in one domain would relax supercoiling in this domain, but not in any of the neighbouring ones (Postow et al. 2004). The domains are not static but dynamic, which is important because replication leads to changes in superhelical density around the forks, which need to travel through the chromo-

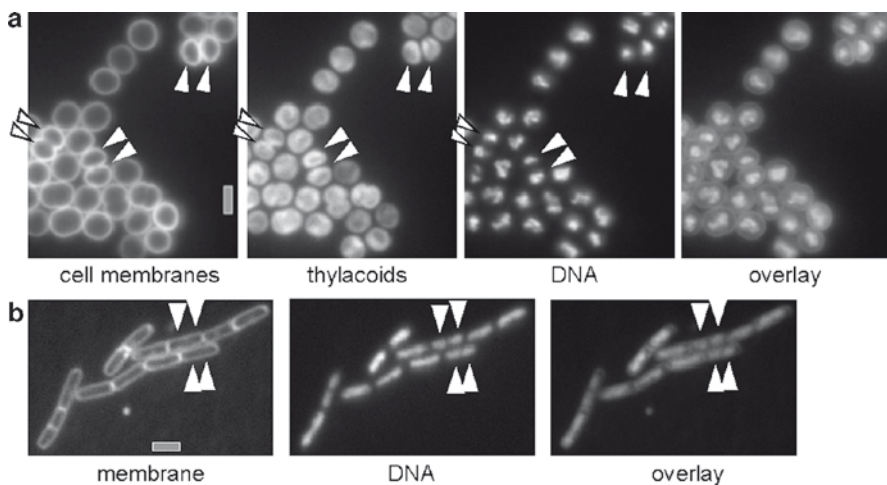


Fig. 3.3 Relaxed chromosome segregation in Cyanobacteria and stringent segregation in *B. subtilis*. (a) Staining of membranes and DNA in *Synechocystis*, arrowheads indicate daughter cells containing different amounts of DNA soon after the division event. (b) In contrast, chromosomes are equally segregated well before a division septum is formed (indicated by arrowheads) in *B. subtilis*. Grey bars 2 μm. Taken from Schneider et al. (2007)

some. However, the molecular basis for the creation of the domains is still unclear. See Chapters 7 and 8 for a discussion of candidate proteins.

Additionally, topoisomerases are essential for chromosome segregation for two main reasons: firstly, replication of helical molecules generates intertwined sister chromosomes behind the replication fork, which need to be unwound, as well as torsional stress in front of the fork (unwinding of the DNA through helices leads to tightening of the double helix in front of the helicase). Gyrase has been shown to be required for ongoing replication, and indeed accumulates at the replication fork in *B. subtilis* cells (Fig. 3.2) (Tadesse and Graumann 2006). Topo IV could also perform this job, but does not accumulate at the replication machinery, and has been found to localize throughout the nucleoid in *B. subtilis* cells. Secondly, after termination of replication, sister chromosomes are still linked, and could not be completely separated, were there not topoisomerases that move the DNA strands of sister chromosomes through each other to allow for complete partitioning. This job is to a major extent performed through Topo IV in *E. coli* (Kato et al. 1990; Zechiedrich et al. 1997), which only becomes active during the last part of the cell cycle (Espeli et al. 2003). Thus, depletion of gyrase or of Topo IV arrests chromosome segregation, due to problems during replication and disentanglement of sister chromosomes. Interestingly, a high concentration of gyrase has been observed in live *B. subtilis* cells close to the cell centre (i.e. close to the replication machinery), whereas Topo I accumulations frequently occur at quarter positions within cells. These differential localization patterns of topoisomerases suggest that regions with different degrees of supercoiling exist on the bacterial nucleoid (Tadesse and Graumann 2006).

3.1.4 SMC

Structural maintenance of chromosomes (SMC) proteins are ubiquitous from bacteria through archaea to eukarya. In eukaryotic cells, they are essential for mitosis and meiosis, and function as central components of several essential protein complexes (cohesin and condensin, as well as DNA repair complexes) (Hirano 2005). SMC proteins exist as dimers, either as heterodimers in eukaryotes, or as homodimers in bacteria. They are composed of a head domain possessing ATPase activity, an extended coiled coil region, and a hinge domain, which mediates specific dimerization. This arrangement sets up a symmetrical molecule with two flexible long arms and two head domains that can also dimerise, dependent on binding of ATP. Dimerization of head domains generates a huge molecular ring that can bind to and thus entrap DNA. Deletion of *smc* in *B. subtilis* or of its ortholog *mukB* in *E. coli* leads to decondensed nucleoids, the production of about 15% cells lacking any DNA (anucleate cells), and to an extreme reduction in growth rate (Britton et al. 1998a; Niki et al. 1992). Most strikingly, the ordered arrangement of genes within the nucleoid is lost in the absence of SMC/MukB (Graumann 2000; Weitao et al. 2000). While SMC is not the motor that drives rapid separation of newly duplicated chromosome regions, its activity is required during each cell cycle, showing that it is a

true segregation factor, and not only functions on global chromosome compaction. In addition to these defects, cells are temperature sensitive, and can grow only up to 23°C in rich medium. In minimal medium, where growth is anyway slower, cells can grow up to 37°C, showing that cells only survive during slow growth, and are killed under conditions of faster growth. BsSMC specifically interacts with ScpA and ScpB, which are widely conserved proteins in prokaryotes, and together, the proteins form a trimeric complex, *in vivo* and *in vitro* (Mascarenhas et al. 2002a, b; Volkov et al. 2003), as do MukB, MukE and MukF (Yamazoe et al. 1999), which are conserved in enterobacteria. Each of these proteins is required for chromosome compaction and segregation, as is SMC or MukB's ATPase activity (Mascarenhas et al. 2005). It is still unclear how the SMC/MukB complexes achieve their tasks in molecular detail, other than that they connect DNA loops with each other in a cooperative and dynamic manner. *In vitro*, SMC and MukB bind tightly to circular DNA, but not to linear DNA (Petrushenko et al. 2006; Volkov et al. 2003), and MukB compacts single DNA molecules in a highly cooperatively manner, generating force-resilient clusters, most likely containing DNA loops (Cui et al. 2008). *In vivo*, SMC/MukB influence supercoiling, because a lack of SMC or of MukB can be compensated for by modulation of Topo I, gyrase or of Topo IV (Lindow et al. 2002a; Sawitzke and Austin 2001; Tadesse et al. 2005). Binding of SMC/MukB to DNA somehow introduces positive writhe into DNA, i.e. the DNA is deformed such that the torsional stress is converted into negative supercoils that e.g. Topo I can relax. Thus, SMC affects global supercoiling in conjunction with topoisomerases. However, different from topoisomerases, SMC acts only at defined positions on the nucleoids. It has clearly been demonstrated that both, SMC and MukB complexes form discrete subcellular assemblies: throughout most of the cell cycle, one assembly can be found within each cell half, away from the replication machinery (Fig. 3.2) (Lindow et al. 2002b; Mascarenhas et al. 2002a; Ohsumi et al. 2001). These assemblies depend on the presence of DNA and on ongoing replication, as well as on proper supercoiling within the cell (Mascarenhas et al. 2005). The following model has been suggested based on several studies: SMC may be loaded onto newly replicated DNA at the replication machinery, and may be translocated towards opposite cell poles with the loops that the complex may induce within the duplicated chromosome regions. SMC/MukB superstructures form within each cell half, in which newly replicated DNA is organized, such that the chromosome obtains its conserved arrangement within each cell cycle (Fig. 3.2). Alternatively, SMC/MukB may form small clusters that dynamically interact with each other to form a large assembly, or even form clusters that dynamically form and dissociate (Cui et al. 2008).

ScpA and ScpB negatively affect SMC's ATPase activity *in vitro* (Kireeva et al. 2004). MukE and MukF also affect MukB's ATPase function (although whether it has a negative or positive effect is the subject of controversy) (Chen and Erickson 2005; Cui et al. 2008), suggesting that the complex partners influence DNA binding of the central SMC subunits. Without ScpA and Scp, SMC loses its specific subcellular localization, which is also the case for mutations in SMC that interfere with ATP binding or ATPase activity. In fact, ATPase mutants of SMC are entirely non-functional *in vivo*. *In vitro*, ScpA and ScpB still bind to ATPase mutants of SMC, but with

lower affinity. These experiments suggest that an interplay of SMC/MukB, ATP, and the two interacting proteins organize DNA binding and compaction of the chromosome. Clearly, SMC/MukB act from the defined subcellular assemblies they form within each cell half, because overproduction of SMC hypercondenses the chromosomes, but remains highly accumulated within the subcellular assemblies within each cell half (Fig. 3.2) (Volkov et al. 2003), in strong contrast to general DNA compaction factors such as HU and H-NS in *E. coli* (Azam et al. 2000) or HBSu in *B. subtilis* that localize throughout the nucleoid (Köhler and Marahiel 1997).

3.1.5 DNA Translocases and Dimer Resolution System

SpoIIIE and FtsK are the founding members of a family of DNA translocases that can form a pore through the membrane and pump DNA using ATP hydrolysis. When DNA is entrapped within the closing division septum (which is always the case during asymmetric division in sporulating *B. subtilis* cells, or can occur after a problem in chromosome segregation during the normal cell cycle), SpoIIIE can translocate DNA into the daughter cells and thus rescue the chromosomes (Wu and Errington 2004). Interestingly, the simultaneous deletion of *smc* and of *spoIIIE* is lethal in *B. subtilis*, showing that *smc* mutant cells can only grow and divide because SpoIIIE rescues non-segregated chromosomes (Britton et al. 1998b). FtsK has a more specialized function in *E. coli*, called dimer resolution (Aussel et al. 2002). Sister chromosomes can recombine with each other, which, for example, can occur right behind the replication forks and does so in a considerable number of growing cells. An uneven number of recombination events will lead to the formation of a chromosome dimer, which cannot be fully segregated (Fig. 3.2, lower left). Bacteria have evolved a system that resolves chromosome dimers in a highly efficient way. An endonuclease/recombinase complex specifically recombines a sequence that is located close to the terminus region on the chromosome, termed *dif* site. This sequence is actively positioned close to the closing division septum by FtsK in *E. coli*, which moves DNA around the terminus region until *dif* sites are in close proximity (Fig. 3.2, lower left panel). In the absence of *dif* or of the recombination system, 10% of the *E. coli* cells show a block in chromosome segregation, demonstrating that dimer formation occurs in a substantial number of cells (Blakely et al. 1993; Kuempel et al. 1991).

3.2 Sloppy Segregation During Growth in Streptomyces, and Stringent Segregation During Differentiation

Streptomyces belong to the high G + C Gram-positive bacteria called actinobacteria, and are a group of bacteria with a unique life style. The model organism *S. coelicolor* grows as branching highly filamentous mycelia (preferentially in the soil), which

extend by polar growth, similarly to fungi. *S. coelicolor* cells divide very irregularly, such that the cells within the filaments contain multiple chromosome copies that are rather randomly positioned (Yang and Losick 2001). In fact, *S. coelicolor* can grow even in the absence of FtsZ, and thus in the absence of cell division. Deletion of SMC or of ParB does not have any effect on cell growth or on nucleoid arrangement within the growing mycelium, suggesting that *Streptomyces* do not need a known active partitioning system during growth (Kim et al. 2000). However, chromosome segregation becomes highly regulated during differentiation. When nutrients become scarce, cells start to grow upwards instead of vertically, and break through the aqueous layer of soil or growth agar. This so-called aerial mycelium (or hyphae) can extend very high into the air, and pinches off many spores at the filament tips in a coordinated and synchronous fashion. Each developing spore compartment must receive a chromosome, and indeed, an active mechanism is likely to be operative, because the error rate of failing chromosome partitioning is extremely low. During this process, SMC and ParB play an important role, because in *smc* or *parB* mutant cells, many spores are devoid of a nucleoid, and because at least ParB assembles into regularly spaced structures within the hyphae (Jakimowicz et al. 2005). Thus, active chromosome segregation does not appear to be required in growing cells, but becomes highly relevant during the developmental process of sporulation, i.e. during special environmental conditions.

3.3 Passive Chromosome Segregation in Cyanobacteria

Many bacteria contain multiple copies of their chromosome, even during slow growth or during their resting state (Breuert et al. 2006). In *Deinococcus radiodurans*, this trait was long thought to confer the extraordinary high resistance to double strand breaks (DSBs) in chromosomal DNA. The *D. radiodurans* chromosomes can be broken more than 100 times (e.g. through gamma irradiation), and are yet efficiently repaired in a few hours. Although repair involves extensive homologous recombination between broken pieces of DNA and extension of ssDNA overhangs along overlapping chromosome fragments by DNA polymerase, radioresistance is also found in *Deinococcus* species that have few or single chromosome copies. Thus, multiple chromosome copies surely facilitate repair of DSBs, but are not a prerequisite for radiation resistance. Rather, the ability to deal with DNA damage through the capturing of DSB-inducing agents such as radicals may be a major underlying mechanism for the high DSB break resistance in *D. radiodurans*.

Possibly, the existence of multiple chromosome copies abolishes the need for an active segregation machinery. Given a large number of chromosomes, it is likely that each daughter cell obtains at least one full copy, by analogy to high copy number plasmids, which lack segregation systems. Such passive chromosome segregation has recently been described for Cyanobacteria, many of which (if not all) contain multiple chromosomes. Two studies have addressed the question of how

multiple chromosomes are separated into daughter cells during the cell cycle in this bacterial phylum. Both studies come to the conclusion that segregation occurs through a more or less random and/or non-regulated mechanism, in which daughter cells receive different numbers of chromosome copies.

Anabaena spp. are filament-forming cyanobacteria. Through the use of the ParB-GFP system, Hu et al. (Hu et al. 2007) have shown that daughter cells receive unequal numbers of chromosome origins. The MreB protein can be lost without a visible effect on chromosome segregation, but is required for the maintenance of proper cell morphology. Thus, clearly, if dedicated segregation proteins exist in *Anabaena*, these do a very sloppy job. In *Synechocystis*, a unicellular, round cyanobacterium, complete chromosome segregation occurs only until very shortly before cell division is terminated, as is apparent from the deeply invaginated dividing cells, in which DNA is still visible within the almost closed septum (Fig. 3.3). Additionally, the intensity of DAPI stained chromosomes between sister cells is very heterogenous: one third/two third segregation patterns are frequently observed (Schneider et al. 2007). In contrast, *B. subtilis* daughter cells usually contain very similar amounts of DNA, as would be expected, and chromosomes are usually well separated before the division septum visibly closes (Fig. 3.2). These two studies strongly suggest that chromosome segregation occurs through a random mechanism in some Cyanobacteria, in which cells divide without complete segregation of chromosomes. How do the cells ensure that those non-segregated chromosomes that are still in the way of the closing septum get translocated? I speculate that FtsK/SpoIIIE-like DNA translocases perform this task and are therefore essential in Cyanobacteria – at least in unicellular species such as *Synechocystis* – because cells would not be able to separate if chromosomes become entrapped in the division septum.

3.4 Massive Polyploidity in a Huge Bacterium

Epulopiscium spp. are among the largest known bacteria, with a size of up to 600 μm . It divides through an interesting mechanism, which involves the production of 2–4 intracellular daughter cells that eventually fill the whole mother cell. Lysis of the mother cell (or what is left) frees the daughter cells. Because *Epulopiscium* is closely related to low G + C Gram-positive bacteria, it is likely that this intracellular production of daughter cells developed from endospore formation. Sporulating *B. subtilis* cells have two chromosomes and produce two FtsZ rings close to each cell pole, one of which forms a septum, while the other dissipates. One chromosome is transported into the forespore, while the other remains in the much larger mother cell. Some mutations lead to the formation of two forespore compartments, which fail to continue to develop, because a mother cell with a full chromosome is required for spore development. During evolution, intracellular forespores could have developed into true daughter cells, however, for this, a mother cell chromosome must have been established in addition to the forespore/daughter cell chromo-

somes. Indeed, *Epulopiscium* has an enormous number of chromosomes, all of which appear to be located just underneath the cell membrane (Mendell et al. 2008). Thus, the daughter cells, which appear to develop through asymmetric division at both cell poles (compared to the mother cell, tiny daughter cells are generated at first), can easily acquire a genome (or even several), while the mother cells retains many genomes. In fact, up to 50,000 genome copies may be present in *Epulopiscium* cells, based on the finding that such a number of gene copies of e.g. *ftsZ* are present in these large cells (Mendell et al. 2008). Thus, *Epulopiscium* has no need for an active chromosome segregation machinery, and can contain a high genetic diversity even within essential genes that is not detrimental to viability. It is also amazing how genetic conservation is achieved by *Epulopiscium* spp. Given the large number of genomes, many mutations will remain silent, but if these happen to be captured in the daughter cell, they may become deleterious. However, only about 1% of the genetic information of the mother cell is transferred to the next generation, because only tiny compartments are divided off (Angert and Clements 2004), so genome stability may be ensured by segregating only a small number of genomes and thus just a few of the mutations that have arisen during the growth cycle.

3.5 Conclusions

It has become clear that bacteria contain dedicated chromosome segregation proteins that ensure high fidelity of partitioning of chromosomes into daughter cells, in addition to the physical properties of DNA that may contribute to chromosome segregation, as discussed in Chapters 5 and 6. In *B. subtilis*, less than 1 in 10,000 cells fails to separate daughter chromosomes properly, a condition where one daughter cell becomes anucleate. Segregation proteins such as SMC are highly conserved from bacteria to man, while ParA type proteins are only present in eubacteria. Many bacteria employ ParA type proteins for plasmid partitioning, while *V. cholerae* appears to use ParA for the segregation of one of its two chromosomes, and a different mechanism – resembling that of *B. subtilis* and of *E. coli* – for the larger chromosome. Segregation in the latter organisms may involve MreB actin-like proteins and a pushing-type mechanism, whereas *C. crescentus* may entirely rely on a ParA type (pulling) segregation mechanism, based on the localization pattern of its chromosome origins and the fact that ParA is essential in this organism. So clearly, several different mechanisms for active chromosome segregation exist in bacteria. It has also become apparent that a non-stringent mode of chromosome segregation exists in some bacteria (e.g. Cyanobacteria) that have multiple chromosome copies, suggesting that bacteria also employ passive and random chromosome segregation based on high copy number. Thus, there is no unified chromosome segregation machinery in bacteria, but apparently several different pathways. Of course, this is not surprising given the high diversity of prokaryotes and their large variation in lifestyles.

Acknowledgements Work in my laboratory is supported by the University of Freiburg and the Deutsche Forschungsgemeinschaft.

References

- Angert ER, Clements KD (2004) Initiation of intracellular offspring in *Epulopiscium*. *Mol Microbiol* 51:827–835
- Aussel L, Barre FX, Aroyo M, Stasiak A, Stasiak AZ, Sherratt D (2002) FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* 108:195–205
- Azam TA, Hiraga S, Ishihama A (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* 5:613–626
- Barilla D, Rosenberg MF, Nobbmann U, Hayes F (2005) Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *EMBO J* 24:1453–1464
- Bates D, Kleckner N (2005) Chromosome and replisome dynamics in *E coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. *Cell* 121:899–911
- Ben-Yehuda S, Losick R (2002) Asymmetric cell division in *B. subtilis* involves a spiral-like intermediate of the cytokinetic protein FtsZ. *Cell* 109:257–266
- Blakely G, May G, McCulloch R, Arciszewska LK, Burke M, Lovett ST, Sherratt DJ (1993) Two related recombinases are required for site-specific recombination at dif and cer in *E coli* K12. *Cell* 75:351–361
- Breuer S, Allers T, Spohn G, Soppa J (2006) Regulated polyploidy in halophilic archaea. *PLoS ONE* 1:e92
- Britton RA, Lin DC-H, Grossman AD (1998a) Characterization of a prokaryotic SMC protein involved in chromosome partitioning. *Genes Dev* 12:1254–1259
- Britton RA, Powell BS, Dasgupta S, Sun Q, Margolin W, Lupski JR, Court DL (1998b) Cell cycle arrest in Era GTPase mutants: a potential growth rate-regulated checkpoint in *Escherichia coli*. *Mol Microbiol* 27:739–750
- Chen Y, Erickson HP (2005) Rapid in vitro assembly dynamics and subunit turnover of FtsZ demonstrated by fluorescence resonance energy transfer. *J Biol Chem* 280:22549–22554
- Cui Y, Petrushenko ZM, Rybenkov VV (2008) MukB acts as a macromolecular clamp in DNA condensation. *Nat Struct Mol Biol* 15:411–418
- Defeu Soufo HJ, Graumann PL (2003) Actin-like proteins MreB and Mbl from *Bacillus subtilis* are required for bipolar positioning of replication origins. *Curr Biol* 13:1916–1920
- Defeu Soufo HJ, Graumann PL (2004) Dynamic movement of actin-like proteins within bacterial cells. *EMBO Rep* 5:789–794
- Defeu Soufo HJ, Graumann PL (2005) *Bacillus subtilis* actin-like protein MreB influences the positioning of the replication machinery and requires membrane proteins MreC/D and other actin-like proteins for proper localization. *BMC Cell Biol* 6:10
- Doi M, Wachi M, Ishino F, Tomioka S, Ito M, Sakagami Y, Suzuki A, Matsushashi M (1988) Determinations of the DNA sequence of the *mreB* gene and of the gene products of the *mre* region that function in formation of the rod shape of *Escherichia coli* cells. *J Bacteriol* 170:4619–4624
- Dworkin J, Losick R (2002) Does RNA polymerase help drive chromosome segregation in bacteria? *Proc Natl Acad Sci USA* 99:14089–14094
- Ebersbach G, Gerdes K (2004) Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Mol Microbiol* 52:385–398
- Ebersbach G, Ringgaard S, Moller-Jensen J, Wang Q, Sherratt DJ, Gerdes K (2006) Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. *Mol Microbiol* 61:1428–1442

- Espeli O, Nurse P, Levine C, Lee C, Marians KJ (2003) SetB: an integral membrane protein that affects chromosome segregation in *Escherichia coli*. *Mol Microbiol* 50:495–509
- Fogel MA, Waldor MK (2005) Distinct segregation dynamics of the two *Vibrio cholerae* chromosomes. *Mol Microbiol* 55:125–136
- Fogel MA, Waldor MK (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev* 20:3269–3282
- Fossum S, Crooke E, Skarstad K (2007) Organization of sister origins and replisomes during multifork DNA replication in *Escherichia coli*. *EMBO J* 26:4514–4522
- Gerdes K, Moller-Jensen J, Ebersbach G, Kruse T, Nordstrom K (2004) Bacterial mitotic machineries. *Cell* 116:359–366
- Gitai Z, Dye NA, Reisenauer A, Wachi M, Shapiro L (2005) MreB actin-mediated segregation of a specific region of a bacterial chromosome. *Cell* 120:329–341
- Gordon S, Sitnikov D, Webb CD, Teleman A, Losick R, Murray AW, Wright A (1997) Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* 90:1113–1121
- Graumann PL (2000) *Bacillus subtilis* SMC is required for proper arrangement of the chromosome and for efficient segregation of replication termini but not for bipolar movement of newly duplicated origin regions. *J Bacteriol* 182:6463–6471
- Hirano T (2005) SMC proteins and chromosome mechanics: from bacteria to humans. *Philos Trans R Soc Lond B Biol Sci* 360:507–514
- Hu B, Yang G, Zhao W, Zhang Y, Zhao J (2007) MreB is important for cell shape but not for chromosome segregation of the filamentous cyanobacterium *Anabaena* sP PCC 7120. *Mol Microbiol* 63:1640–1652
- Jacob F, Brenner S, Cuzin F (1963) On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp Quant Biol* 28:329–348
- Jakimowicz D, Gust B, Zakrzewska-Czerwinska J, Chater KF (2005) Developmental-stage-specific assembly of ParB complexes in *Streptomyces coelicolor* hyphae. *J Bacteriol* 187:3572–3580
- Jensen SO, Thompson LS, Harry EJ (2005) Cell division in *Bacillus subtilis*: FtsZ and FtsA association is Z-ring independent, and FtsA is required for efficient midcell Z-Ring assembly. *J Bacteriol* 187:6536–6544
- Jones LJ, Carballido-Lopez R, Errington J (2001) Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. *Cell* 104:913–922
- Kato J, Nishimura Y, Imamura R, Niki H, Hiraga S, Suzuki H (1990) New topoisomerase essential for chromosome segregation in *E coli*. *Cell* 63:393–404
- Kim HJ, Calcutt MJ, Schmidt FJ, Chater KF (2000) Partitioning of the linear chromosome during sporulation of *Streptomyces coelicolor* A3(2) involves an *oriC*-linked *parAB* locus. *J Bacteriol* 182:1313–1320
- Kireeva N, Lakonishok M, Kireev I, Hirano T, Belmont AS (2004) Visualization of early chromosome condensation: a hierarchical folding, axial glue model of chromosome structure. *J Cell Biol* 166:775–785
- Köhler P, Marahiel MA (1997) Association of the histone-like protein HBSu with the nucleoid of *Bacillus subtilis*. *J Bacteriol* 179:2060–2064
- Kruse T, Moller-Jensen J, Lobner-Olesen A, Gerdes K (2003) Dysfunctional MreB inhibits chromosome segregation in *Escherichia coli*. *EMBO J* 22:5283–5292
- Kruse T, Blagoev B, Lobner-Olesen A, Wachi M, Sasaki K, Iwai N, Mann M, Gerdes K (2006) Actin homolog MreB and RNA polymerase interact and are both required for chromosome segregation in *Escherichia coli*. *Genes Dev* 20:113–124
- Kuempel PL, Henson JM, Dircks L, Tecklenburg M, Lim DF (1991) *dif*, a *recA*-independent recombination site in the terminus region of the chromosome of *Escherichia coli*. *New Biol* 3:799–811
- Lemon KP, Grossman AD (1998) Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282:1516–1519
- Lindow JC, Britton RA, Grossman AD (2002a) Structural maintenance of chromosomes protein of *Bacillus subtilis* affects supercoiling *in vivo*. *J Bacteriol* 184:5317–5322

- Lindow JC, Kuwano M, Moriya S, Grossman AD (2002b) Subcellular localization of the *Bacillus subtilis* structural maintenance of chromosomes (SMC) protein. *Mol Microbiol* 46:997–1009
- Margolin W (2005) FtsZ and the division of prokaryotic cells and organelles. *Nat Rev Mol Cell Biol* 6:862–871
- Mascarenhas J, Soppa J, Strunnikov A, Graumann PL (2002a) Cell cycle dependent localization of two novel prokaryotic chromosome segregation and condensation proteins in *Bacillus subtilis* that interact with SMC protein. *EMBO J* 21:3108–3118
- Mascarenhas J, Soppa J, Strunnikov AV, Graumann PL (2002b) Cell cycle-dependent localization of two novel prokaryotic chromosome segregation and condensation proteins in *Bacillus subtilis* that interact with SMC protein. *EMBO J* 21:3108–3118
- Mascarenhas J, Volkov AV, Rinn C, Schiener J, Guckenberger R, Graumann PL (2005) Dynamic assembly, localization and proteolysis of the *Bacillus subtilis* SMC complex. *BMC Cell Biol* 6:28
- Mendell JE, Clements KD, Choat JH, Angert ER (2008) Extreme polyploidy in a large bacterium. *Proc Natl Acad Sci USA* 105:6730–6734
- Niki H, Imamura R, Kitaoka M, Yamanaka K, Ogura T, Hiraga S (1992) *E coli* MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. *EMBO J* 11:5101–5109
- Niki H, Yamaichi Y, Hiraga S (2000) Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14:212–223
- Ohsumi K, Yamazoe M, Hiraga S (2001) Different localization of SeqA-bound nascent DNA clusters and MukF-MukE- MukB complex in *Escherichia coli* cells. *Mol Microbiol* 40:835–845
- Peters PC, Migocki MD, Thoni C, Harry EJ (2007) A new assembly pathway for the cytokinetic Z ring from a dynamic helical structure in vegetatively growing cells of *Bacillus subtilis*. *Mol Microbiol* 64:487–499
- Petrushenko ZM, Lai CH, Rybenkov VV (2006) Antagonistic interactions of kleisins and DNA with bacterial Condensin MukB. *J Biol Chem* 281:34208–34217
- Postow L, Hardy CD, Arsuaga J, Cozzarelli NR (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev* 18:1766–1779
- Reyes-Lamothe R, Possoz C, Danilova O, Sherratt DJ (2008) Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* 133:90–102
- Sawitzke J, Austin S (2001) An analysis of the factory model for chromosome replication and segregation in bacteria. *Mol Microbiol* 40:786–794
- Schneider D, Fuhrmann E, Scholz I, Hess WR, Graumann PL (2007) Fluorescence staining of live cyanobacterial cells suggest non-stringent chromosome segregation and absence of a connection between cytoplasmic and thylakoid membranes. *BMC Cell Biol* 8:39
- Srivastava P, Demarre G, Karpova TS, McNally J, Chatteraj DK (2007) Changes in nucleoid morphology and origin localization upon inhibition or alteration of the actin homolog, MreB, of *Vibrio cholerae*. *J Bacteriol* 189:7450–7463
- Stacek P, Higgins NP (1998) Gyrase and Topo IV modulate chromosome domain size in vivo. *Mol Microbiol* 29:1435–1448
- Stone MD, Bryant Z, Crisona NJ, Smith SB, Vologodskii A, Bustamante C, Cozzarelli NR (2003) Chirality sensing by *Escherichia coli* topoisomerase IV and the mechanism of type II topoisomerases. *Proc Natl Acad Sci USA* 100:8654–8659
- Sunako Y, Onogi T, Hiraga S (2001) Sister chromosome cohesion of *Escherichia coli*. *Mol Microbiol* 42:1233–1242
- Tadesse S, Graumann PL (2006) Differential and dynamic localization of topoisomerases in *Bacillus subtilis*. *J Bacteriol* 188:3002–3011
- Tadesse S, Mascarenhas J, Kusters B, Hasilik A, Graumann PL (2005) Genetic interaction of the SMC complex with topoisomerase IV in *Bacillus subtilis*. *Microbiology* 151:3729–3737
- Teleman AA, Graumann PL, Lin DCH, Grossman AD, Losick R (1998) Chromosome arrangement within a bacterium. *Curr Biol* 8:1102–1109
- Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Natl Acad Sci USA* 101:9257–9262

- Volkov AV, Mascarenhas J, Andrei-Selmer C, Ulrich HD, Grauman PL (2003) A prokaryotic condensin/cohesin-like complex can actively compact chromosomes from a single position on the nucleoid. *Mol Cell Biol* 23:5638–5650
- Wang JC, (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol* 3:430–440
- Webb CD, Teleman A, Gordon S, Straight A, Belmont A, Lin DC-H, Grossman AD, Wright A, Losick R (1997) Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* 88:667–674
- Webb CD, Graumann PL, Kahana J, Teleman AA, Silver P, Losick R (1998) Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28:883–892
- Weitao T, Dasgupta S, Nordstrom K (2000) Role of the *mukB* gene in chromosome and plasmid partition in *Escherichia coli*. *Mol Microbiol* 38:392–400
- Woldringh CL, Nanninga N (2006) Structural and physical aspects of bacterial chromosome segregation. *J Struct Biol* 156:273–283
- Wu LJ, Errington J (2004) Coordination of cell division and chromosome segregation by a nucleoid occlusion protein in *Bacillus subtilis*. *Cell* 117:915–925
- Yamazoe M, Onogi T, Sunako Y, Niki H, Yamanaka K, Ichimura T, Hiraga S (1999) Complex formation of MukB, Muke and MukF proteins involved in chromosome partitioning in *Escherichia coli*. *EMBO J* 18:5873–5884
- Yang MC, Losick R (2001) Cytological evidence for association of the ends of the linear chromosome in *Streptomyces coelicolor*. *J Bacteriol* 183:5180–5186
- Zechiedrich EL, Khodursky AB, Cozzarelli NR (1997) Topoisomerase IV, not gyrase, decatenates products of site-specific recombination in *Escherichia coli*. *Genes Dev* 11:2580–2592
- Zechiedrich EL, Khodursky AB, Bachellier S, Schneider R, Chen D, Lilley DM, Cozzarelli NR (2000) Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli*. *J Biol Chem* 275:8103–8113

Chapter 4

Extrachromosomal Components of the Nucleoid: Recent Developments in Deciphering the Molecular Basis of Plasmid Segregation

Finbarr Hayes and Daniela Barillà

Abstract Plasmids are extrachromosomal elements that are widely distributed in eubacteria, as well as in archaea and lower eukaryotes. Plasmids confer additional genetic plasticity on species that harbour them, but also are of major clinical significance because antibiotic resistance, virulence, and other disease-associated genes often reside on these highly mobile elements. Moreover, plasmids are malleable and informative models to improve understanding of bacterial genome segregation: the molecular mechanisms of bacterial DNA segregation are best described for low copy number plasmids. The segrosome is the nucleoprotein complex that drives accurate plasmid partitioning. The complex typically includes: (i) a centromere analogue on which segrosome assembly occurs; (ii) one of a diverse array of site-specific DNA binding factors that recognizes its cognate centromere and with which it forms a nucleoprotein structure of specific architecture; and (iii) an ATP binding protein, either actin-like or, more commonly, a Walker-type ATPase of the ParA superfamily that is unique to prokaryotes and which assembles into the mature segrosome. ATP-mediated polymerization of actin-like segregation proteins into a bipolar spindle elicits bidirectional filament growth, propelling attached plasmids in opposing directions prior to cytokinesis. Plasmid-encoded ParA proteins also polymerize in response to ATP binding, although the molecular mechanisms that underpin this behaviour and how this polymerization mediates intracellular plasmid trafficking remain to be fully elucidated. Recent insightful biochemical, structural and cell biological analyses of segrosome assembly and action continue to unravel fundamental aspects of plasmid segregation.

Keywords Plasmid • segregation • partition • ParA/actin

F. Hayes (✉)

Faculty of Life Sciences and Manchester Interdisciplinary Biocentre, The University of Manchester, 131 Princess Street, Manchester M1 7DN, UK
e-mail: finbarr.hayes@manchester.ac.uk

D. Barillà

Department of Biology, University of York, P.O. Box 373, York YO10 5YW, UK
e-mail: db530@york.ac.uk

4.1 Introduction

Plasmids are extrachromosomal elements that are widely prevalent in diverse bacterial species, as well as more restrictedly in archaea and lower eukaryotes (Hayes 2003a). Plasmids vary considerably in size from cryptic elements <2 kb in length that possess only sufficient information for their replication, to megaplasmids many hundreds of kilobases in size that may constitute a significant fraction of the host genome. Accessory genes located on certain plasmids permit their hosts to proliferate in environmental niches that they might not otherwise be able to occupy. These accessory functions can prove to be of major clinical concern in the case of plasmids that specify resistance to one or multiple antibiotics, that encode toxins, or which promote bacterial virulence or pathogenesis. These concerns are exacerbated by plasmids' innate ability to acquire new genetic traits by recombination or transposition, as well as to spread rapidly in bacterial populations by horizontal transfer (Thomas 2000).

A characteristic feature of naturally occurring plasmids is that they are maintained at distinctive copy numbers under steady state conditions. Plasmid copy number is dictated by control circuits which modulate the replication frequency in tune with cell growth rate and other physiological fluctuations (Chattoraj 2000; del Solar and Espinosa 2000; Nordstrom 2006). Although some plasmids are present at tens of copies per cell, many plasmids are available at just two or a few copies during cytokinesis. If these low copy number plasmids relied solely on passive cytoplasmic diffusion to ensure their accurate transmission to daughter cells, the emergence of plasmid-free cells would be sufficiently frequent that maintenance of the plasmid in the bacterial population would be jeopardized. In fact, low copy number plasmids are inherited with remarkable fidelity, even in the absence of selective pressure, revealing that they harbour dedicated mechanisms that guarantee their faithful distribution to new cells. Plasmids often mediate site-specific recombination reactions that convert to monomers any plasmid dimers or multimers that arise by homologous recombination (Summers 1998). This process optimizes the number of plasmid units available for distribution at cell division. Plasmids also commonly specify toxin-antitoxin complexes that kill plasmid-free cells postsegregationally (Hayes 2003b). The antitoxin is more susceptible to host proteases than is the toxin. When a plasmidless cell arises, the toxin is released from its association with the depleted antitoxin. As the antitoxin cannot be replenished in the plasmid-free cell, the toxin is available to target an essential intracellular host component to induce cell death or severe growth restriction.

4.2 Plasmid Segregation: Four Types of Modules

Additional to multimer resolution and postsegregational killing mechanisms, low copy number plasmids typically ensure their active partition to daughter cells using a dedicated segregation locus. There has been a wealth of genetic, biochemical and cell biological analysis of the molecular mechanisms that underpin formation of the

segrosome, the nucleoprotein complex that directs plasmid partition (Hayes and Barillà 2006a, b), which has provided key insights into the process. Segregation cassettes have been classified into four types (I–IV) based on the genetic organization of the module and the evolutionary relationships among the encoded proteins (Fig. 4.1) (Ebersbach and Gerdes 2005a; Hayes and Barillà 2006a, b; Schumacher 2007, 2008). The type I partition system is the most prevalent and, along with the type II complex, has been most extensively studied. The most well-characterized

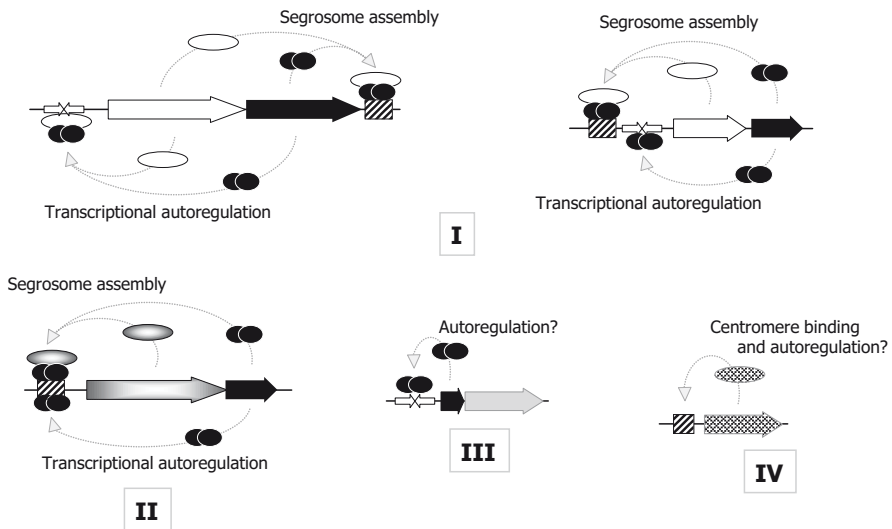


Fig. 4.1 Genetic organization and interactions in plasmid segregation cassettes. The widely-disseminated **Type I** partition modules possess genes for a ParA Walker-type ATPase and a CBF. The genes for *parA* and the CBF, and the corresponding proteins are shown by open and filled symbols, respectively. The centromere located either upstream or downstream of the genes is indicated by a hatched box. The operator site upstream of the genes is shown schematically by small, inverted arrows. During segrosome formation, the CBF binds to the centromere and recruits the ParA factor which itself does not bind centromeric DNA. In some cases, the ParA protein is the principal transcriptional autoregulator of the *par* operon, whereas the ParB CBF is a co-repressor that does not interact directly with the operator site (*top left*). In other cases, the CBF binds to both the centromere and regulatory region (*top right*). In these cases, the ATPase is recruited to the segrosome via protein-protein interactions with the CBF, but does not participate in transcriptional autoregulation. In the well-characterized **type II** module, a pair of genes specify an actin-like ATPase (ParM; *shaded ovoid*), and a small DNA binding protein (ParR). The centromere and operator site located upstream of the genes overlap. ParR is both a CBF, as well as a transcriptional autorepressor. ParM interacts with ParR at the centromere, but is not involved in gene regulation. **Type III** partition cassettes are found on certain large plasmids of the *Bacillus cereus* group and comprise two genes. The first gene encodes the RepR protein that is a likely transcriptional autorepressor, whereas the following gene (*grey arrow*) encodes a member of the tubulin family known as TubZ or RepX. **Type IV** segregation modules include a single gene (*hatched arrow*) whose product is predicted to possess both DNA binding properties, as well as a region of coiled-coil structure

examples of these two types are discussed here first, with a subsequent brief description of type III and IV classes.

Type I cassettes are the most widespread segregation modules, and are found very commonly on plasmids in a wide diversity of eubacteria, as well as in archaea. The cassettes are disparate in composition and organization, but are unified by the presence of genes specifying Walker-type ATPases that have a close evolutionary origin (Fig. 4.2). These motor-like ATPases, commonly denoted ParA, possess divergent Walker-like ATP binding motifs (Koonin 1993; Motallebi-Veshareh et al. 1990). The ParA proteins range in size from ~200 amino acids to ~450 amino acids. Type I modules invariably contain a second gene downstream of, and transcribed in the same direction as, the gene for ParA. The encoded product is a DNA binding factor that recognizes the *cis*-acting centromere sequence located within the cassette, forming a pre-segrosome complex. These centromere binding factors (CBFs) are more diverse than the ParA components, both in size and sequence (Fothergill et al. 2005; Hayes and Barillà 2006a). However, one subset of CBFs, exemplified by ParB of plasmid P1, is particularly common and is frequently specified by type I cassettes that encode large ParA proteins. Many bacterial chromosomes also harbour *parAB* cassettes (Bartosik and Jagura-Burdzy 2005; Thanbichler and Shapiro 2006).

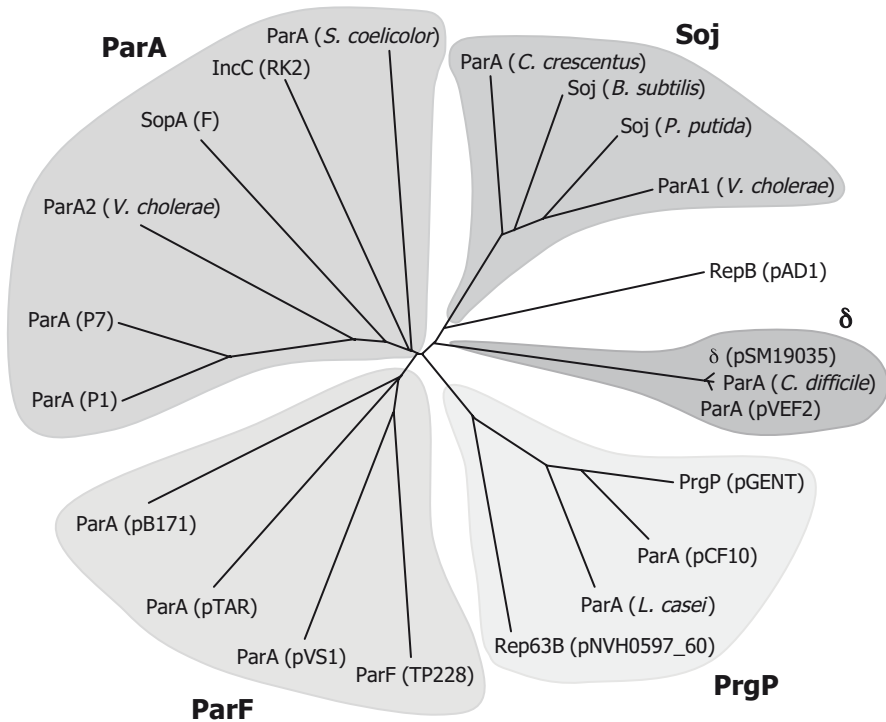


Fig. 4.2 Phylogenetic tree of selected ParA family members. Plasmid-encoded proteins only are shown, except for the Soj group of chromosomally-encoded homologues, and the ParA and ParA2 proteins specified by the *S. coelicolor* chromosome and chromosome II of *V. cholerae*, respectively

By contrast, plasmid segregation modules that possess the shortest *parA* loci include genes for CBFs that may have very little, if any, amino acid sequence similarity and which preferentially recognize the cognate centromeric sites (Fothergill et al. 2005). Nevertheless, two of these disparate CBFs, ParG and ω of plasmids TP228 and pSM19035, respectively, are homodimers that possess ribbon-helix-helix (RHH) folds which bind DNA (Golovanov et al. 2003; Murayama et al. 2001), hinting that common structural motifs may be employed for centromere recognition among smaller CBFs within the type I class.

Type II modules, typified by the partition cassette of plasmid R1, have a similar genetic organization to type I operons (Fig. 4.1). Type II modules comprise two genes, one of which encodes an ATPase (ParM) and the second of which specifies a CBF (ParR). The centromere (*parC*) is situated upstream of the genes (Salje and Lowe 2008). However, by contrast with type I cassettes, the motor-like ATPase specified by type II systems is an ancestral homologue of eukaryotic actin (Bork et al. 1992), and is unrelated evolutionarily to ParA proteins. A series of elegant studies, described further in Section 4.5, have illuminated core aspects of the type II partitioning mechanism.

Both type I and II modules are transcriptionally autoregulated (Fig. 4.1). Large ParA homologues repress transcription via an N-terminal helix-turn-helix (HTH) motif that binds to an operator site located 5' of the partition cassette. The CBF is a co-repressor that decreases operon expression further by an unknown mechanism. Autoregulation of the *parAB* operon of the P1 plasmid arguably is the most well-studied example among segregation cassettes (Davey and Funnell 1994; Friedman and Austin 1988; Hayes et al. 1994; Radnedge et al. 1998). The nucleotide bound state of ParA modulates transcriptional regulation: ADP, but not ATP, promotes ParA repression of the operon (Bouet and Funnell 1999). ADP may induce a conformational change in the protein so that it is more proficient than nucleotide-free ParA in DNA binding. Instead, the ATP-bound form of ParA is required for segregation (Davis et al. 1996). Most plausibly, ATP fulfils an analogous function in P1 segregation as it does in other type I complexes, by promoting ParA polymerization (see Section 4.4). By contrast with P1 ParA and its immediate homologues, both short ParA type I proteins and the ParM type II ATPase lack DNA-binding motifs and are not implicated in transcriptional regulation of their operons. Instead, the CBFs alone are transcriptional repressors of the operons, as well as integral components of the segrosome (Fig. 4.1) (e.g., Carmelo et al. 2005; Dmowski et al. 2006; Jensen et al. 1994; Ringaard et al. 2007a; Zampini et al. 2009).

4.3 Centromere Binding Factors: Underpinning the Segrosome

Centromeres in higher eukaryotes are DNA regions that direct kinetochore formation and the cohesion of sister chromatids. The structural features of eukaryotic centromeres, which in some species extend for many thousands of kilobases, continue to defy full understanding. Like centromeres in eukaryotes, plasmid centromeres

are remarkably diverse in sequence, in length, and in the numbers, orientation and spacing of repeated sequences that they contain (Hayes and Barillà 2006a) (Fig. 4.3). Nevertheless, plasmid centromeres typically are ~100–200 bp in length

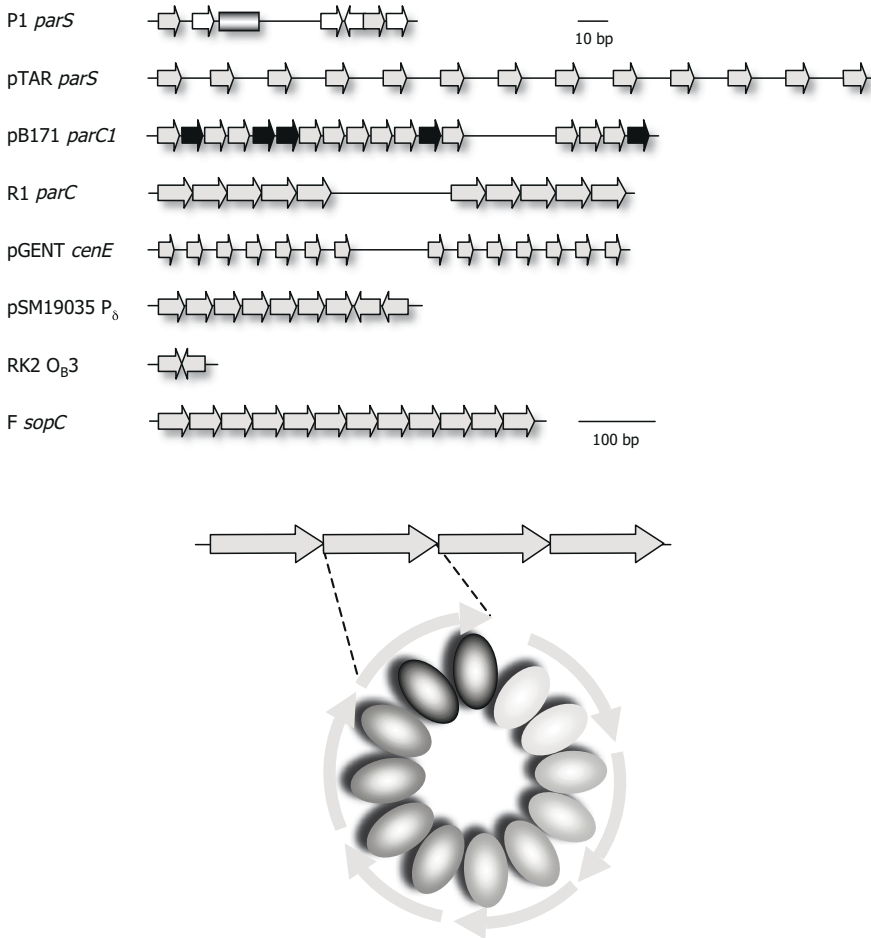


Fig. 4.3 Plasmid centromere organization. *Top*, the orientation of repeat motifs in selected centromeres is shown by arrows. The lengths and sequences of these motifs differ between different centromeres. Additionally, the P1 *parS* and pB171 *parC1* sites each possess two different repeat sequences highlighted by different arrow shadings. The binding site for IHF in P1 *parS* is denoted by the rectangle. The P_{δ} region is one of three loci on pSM19035 that possess centromere activity. Similarly, a second set of repeat motifs in pB171 is implicated in centromere function. A 10-bp scale bar is shown, except for *F sopC* for which a 100-bp scale is shown. *Bottom*, schematic representation of co-crystal structures of ParR dimers (oval) of plasmid pSK41 bound to one repeat motif (arrow) from its cognate centromere (Schumacher et al. 2007a). A pair of ParR_{SK41} dimers binds to a single repeat. Interactions between dimers bound to separate DNA fragments mediate the assembly of a nucleoprotein superstructure containing a ‘pseudo-centromere’-like sequence wrapped around a ParR_{SK41} protein core. The ParR protein of pSK41 is homologous to the ParR proteins encoded by plasmids R1 and pB171. Dimers of the latter also coalesce into ring-like structures both in the absence and presence of centromere DNA (Moller-Jensen et al. 2007)

(Fig. 4.3), and occupy defined locations either upstream or downstream of the partition genes (Fig. 4.1). Despite their diversity, centromeres uniformly provide a scaffold for loading of the cognate CBF (Fig. 4.4a). Elucidation of the tertiary structures of a number of CBFs, in some instances complexed with centromeric DNA within the pre-segrosomal complex, recently have begun to provide fascinating glimpses into assembly of the mature segrosome (de la Hoz et al. 2004; Delbrück et al. 2002; Golovanov et al. 2003; Khare et al. 2004; Moller-Jensen et al. 2007; Murayama et al. 2001; Schumacher 2007, 2008; Schumacher et al. 2007b).

The *parS* site of P1 is among the most well-studied plasmid centromeres. The site comprises two sets of repeat motifs separated by a central region that is bound by the DNA bending protein, integration host factor (IHF). The repeats are of two distinct types, either Box A (heptameric) or Box B (hexameric) (Fig. 4.3). The dimeric ParB CBF recognizes the Box A and Box B motifs using separate protein domains. Amino terminal domains of ParB recognize the Box A repeats via HTH motifs, whereas a dimerized DNA-binding module composed of a six-stranded β -sheet coiled-coil interacts with the hexamer boxes near to the ends of the site (Schumacher and Funnell 2005; Vecchiarelli et al. 2007). A HTH motif has also been implicated in DNA binding by the ParB homologue, KorB, of plasmid RK2 (Khare et al. 2004). The DNA-binding domains within ParB are connected by flexible linkers around which the domains can swivel freely to permit multiple arrangements of ParB-*parS* contacts. Indeed, the ParB dimer simultaneously can make in trans contacts with *parS* sites on different DNA molecules (Schumacher and Funnell 2005; Schumacher et al. 2007a), supporting observations that ParB can pair plasmids that harbour the site (Edgar et al. 2001). The role of IHF is entirely architectural, serving to bend the arms of *parS* to accommodate ParB dimers that span the two arms (Funnell 1988, 1991; Hayes and Austin 1994). The *parS* locus is a nucleation point for the spreading of ParB many kilobases away from the site (Rodionov et al. 1999), as is the *sopC* centromere for the ParB homologue encoded by the F plasmid (Lynch and Wang 1995). It is tempting to speculate that this reflects the formation of a nucleoprotein superstructure in the vicinity of the centromere, but the purpose, if any, of this spreading during segregation remains unresolved (Rodionov and Yarmolinsky 2004).

Variations in the Box B repeats in the *parS* site and corresponding substitutions in the dimerized DNA-binding module of ParB provide species specificity among segrosomes that are closely-related to the P1 partition complex (Hayes and Austin 1993; Hayes et al. 1993; Radnedge et al. 1996, 1998; Sergueev et al. 2005). Indeed, a single alteration in the Box B repeats and a matching substitution in the one of the ParB residues that contacts this position can induce a comprehensive switch in the specificity of the interactions (Dabrazhynetskaya et al. 2005, 2009). These natural variations potentially permit plasmids with closely-related partition modules to co-exist without cross-interference in segrosome assembly. An alternative evolutionary strategy to this end is the acquisition by plasmids of CBFs that recognize centromeres which differ majorly in the number, arrangement and sequences of the repeat motifs that they contain. The resulting CBF-centromere interactions are exclusive, conferring a layer of specificity to the macromolecular interactions that mediate plasmid segregation (Fothergill et al. 2005).

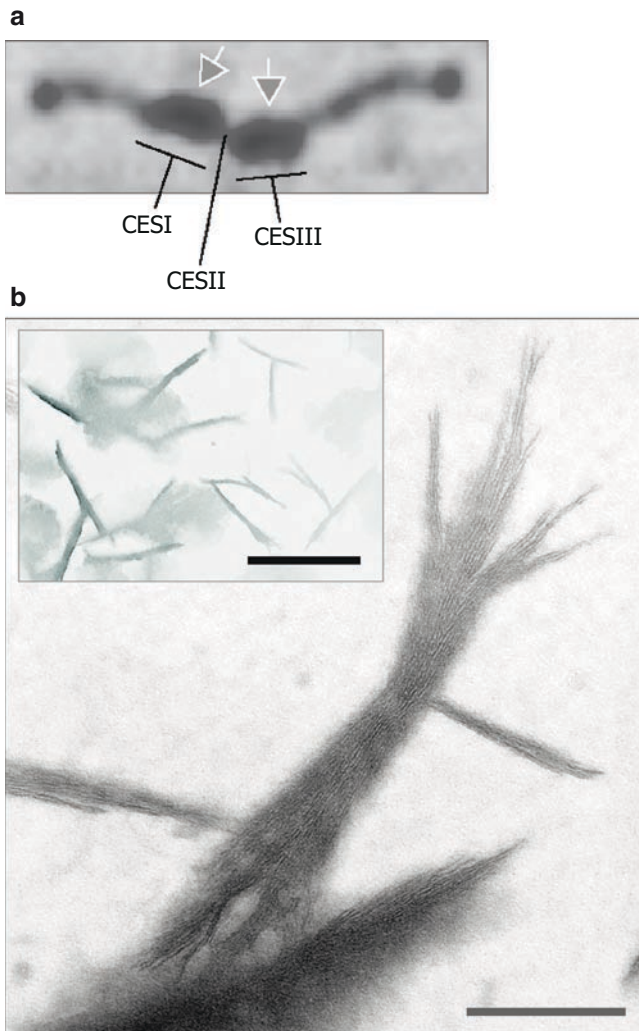


Fig. 4.4 (a) Atomic force microscopy of the PrgP CBF of plasmid pGENT loaded on the *cenE* centromere located within a linear DNA fragment (Derome et al. 2008; M. Bussiek, A. Derome, C. Hoischen, S. Diekmann, D. Barillà, and F. Hayes, unpublished data). PrgP recognizes two arrays of seven TATA boxes (CESI and CESIII) separated by the CESII spacer in *cenE* (Fig. 4.3). The PrgP foci at these two arrays are highlighted with arrows. (b) Electron microscopy reverse contrast images of multistranded ParF polymers (*inset*) and ParF fibres assembled in the presence of ParG (*main*). Bar = 210 nm (*main*) and 500 nm (*inset*)

The ParR protein of plasmid R1 binds to two arrays of five 11-bp direct repeats in the *parC* centromere that sandwich a spacer region containing the *parMR* promoter (Fig. 4.3) (Breüner et al. 1996; Dam and Gerdes 1994; Hoischen et al. 2008; Salje and Lowe 2008). Whereas IHF bends the *parS* site of the P1 plasmid, a

combination of intrinsic curvature and ParR-induced bending severely distorts *parC* into a U-shaped structure (Hoischen et al. 2004, 2008). ParR of plasmid R1 has proven refractory to detailed structural studies, but dimeric sequence homologues encoded by pB171 of *E. coli* and staphylococcal plasmid pSK14 possess RHH folds that mediate binding to their cognate centromeres. Intriguingly, the crystal structure of ParR_{B171} is arranged into a continuous helical assembly comprising 12 dimers per 360° turn. These closed rings are also visible by electron microscopy in the presence of centromere DNA (Moller-Jensen et al. 2007). Analogously, pairs of ParR_{SK41} dimers bound to a minimal centromere subsite coalesce into an extended looped structure held in place by protein-protein interactions (Fig. 4.3) (Schumacher et al. 2007a). The DNA binding domains of both ParR_{B171} and ParR_{SK41} within these ring-shaped superstructures point away from the central cavity, suggesting that the centromeric DNA wraps around the external surface of the rings (Moller-Jensen et al. 2007; Schumacher et al. 2007a). It has been hypothesized that the extended ParM filament (see Section 4.5) may be tethered within the central cavity of the ParR ring in the mature segrosome (Moller-Jensen et al. 2007).

Like yeast centromeres, the *cenE* site of enterococcal plasmid pGENT is intrinsically curved, albeit less than the R1 *parC* site (Derome et al. 2008). The *cenE* centromere comprises three subregions: the CESI and CESIII subsites are separated by the CESII spacer (Fig. 4.3). CESI and CESIII each contain seven TATA boxes spaced by half-helical turns. The PrgO CBF independently binds CESI and CESIII, but with different avidities, whereas the protein does not occupy the CESII subsite. The function of the P1 *parS* site is unimpaired by the insertion of integral helical turns between the site's arms (Hayes and Austin 1994), reflecting the relative flexibility of the DNA binding domains in the ParB dimers that span the arms (Schumacher and Funnell 2005). However, *parS* activity is abolished by the insertion of non-integral helical turns in the centre of the site, indicating that the arms must be positioned with specific faces of the two helices facing each other (Hayes and Austin 1994). By contrast, *cenE* is tolerant to insertions of both integral and non-integral turns in the CESII subsite suggesting that the architecture of the PrgO-*cenE* pre-segrosome complex differs from that of ParB-*parS* (Derome et al. 2008).

Like ParR, the ω CBF possesses a RHH fold that binds DNA (Murayama et al. 2001). The ω protein was first established as a global transcriptional repressor of pSM19035 genes (de la Hoz et al. 2000), before its role as a CBF was clarified more recently (Dmowski et al. 2006; Pratto et al. 2008). The protein recognizes multiple heptad motifs in its binding sites; at least three of these sites may act as centromeric sequences (Fig. 4.3). Co-crystal structures and footprinting data of ω on these sites suggest an elongated protein superstructure in which ω dimers spiral as a left-handed helix that enwraps the centromeric DNA (de la Hoz et al. 2004; Weihofen et al. 2006). Thus, an emerging theme in the interaction of CBFs with their binding sites is the formation of intricate nucleoprotein superstructures that underpin the segrosome. The specific conformation of these structures is likely to be crucial for interaction with the motor-like ParA or ParM polymeric protein and the assembly of the functional segrosome.

4.4 Plasmid-Encoded ParA Proteins are Polymerizing Motor-Like ATPases

Although it has been recognized for some time that ParA proteins encoded by type I modules possess Walker box ATP binding and hydrolysis motifs (Koonin 1993; Motallebi-Veshareh et al. 1990), and that disruption of these motifs abolished plasmid partitioning in vivo (Davis et al. 1996), the purpose of nucleotide binding and cleavage was uncertain. Was ATP hydrolysis the driving force for plasmid movement? If so, how might the comparatively weak ATPase activity of ParA proteins provide sufficient energy to propel plasmids through the dense cytoplasm? Recent studies demonstrating that plasmid-encoded ParA proteins polymerize in vivo and in vitro in response to ATP binding have begun to answer these conundrums (Adachi et al. 2006; Barillà et al. 2005, 2007; Bouet et al. 2007; Ebersbach and Gerdes 2001, 2005b; Machón et al. 2007; Pratto et al. 2008). The ParA homologue, ParF, encoded by the TP228 multiresistance plasmid (Barillà and Hayes 2003; Hayes 2000) assembles into extensive multistranded filaments in response to ATP binding in vitro. Nucleotide hydrolysis is not required for filamentation. Moreover, mutations within the Walker box motifs of ParF perturb polymerization, and ADP blocks polymerization, confirming the role of nucleotide binding in filamentation (Barillà et al. 2005). ParF polymerization is modulated by the ParG CBF, both in the presence and absence of nucleotide (Fig. 4.4b). More specifically, an unstructured N-terminal region of ParG is required for stimulation of ParF polymerization (Barillà et al. 2007). Although the mechanism by which the ParG flexible tail affects ParF polymerization has yet to be revealed, it might act analogously to the mobile ‘tentacles’ of actin capping protein that regulate polymerization by binding the barbed ends of actin filaments (Cooper and Sept 2008). Interestingly, CBFs with unrelated primary sequences can also promote ParF polymerization in vitro suggesting that a common mechanism may exist by which diverse CBFs modulate the polymerization of homologous ParA proteins (Machón et al. 2007). Elucidation of the tertiary structures of additional CBFs will reveal whether they also possess flexible regions that interact with ParA proteins and modulate their polymerization kinetics.

In common with numerous other bacterial cytoskeletal proteins (Pogliano 2008), the ParA homologues of plasmids pB171 (ParA_{B171}), F (SopA) and pSM19035 (δ) display distinct subcellular localization patterns, as well as polymerizing in vitro in response to ATP binding. ParA_{B171} oscillates along a spiral structure that is distributed over the nucleoid. Oscillation requires both other components of the segregationosome, the ParB_{B171} CBF and the cognate centromere, *parC*, as well as intact Walker box motifs within ParA_{B171}. The three elements coordinately position plasmids at regular intervals along the main axis of the cell (Ebersbach and Gerdes 2001; Ebersbach et al. 2006). The oscillation of ParA_{B171} mimics the behaviour of MinD proteins that form a discrete branch of the ParA superfamily. MinD, in concert with MinC, inhibits random placement of the bacterial cell division septum. Inhibition is relaxed specifically at the cell centre by the MinE factor, thereby permitting assembly of the cell division apparatus only at the correct location

(Rothfield et al. 2005). Nevertheless, despite their common evolutionary origin and related subcellular localization patterns, ParA_{B171} appears to function independently of MinD, as well as of other components of the cell division complex (Ebersbach and Gerdes 2001).

Like ParA_{B171}, the ParA homologue (SopA) encoded by the F plasmid oscillates between the one- and three-quarter cell length positions in *E. coli* with a periodicity of ~20 min. This localization pattern requires the SopB CBF and the *sopC* centromere and likely reflects, first, the formation of SopA helical structures nucleated by SopB loaded at *sopC*, followed by the disassembly of the SopA polymers (Adachi et al. 2006; Hatano et al. 2007; Lim et al. 2005). However, SopA filaments also have been reported to form independently of SopB. Moreover, SopB also aggregates into helical structures which partially overlap with SopA spirals, and which require SopA for their formation in vivo (Adachi et al. 2006). SopA also polymerizes in vitro into filaments with ultrastructures that resemble those of ParF, and whose formation is influenced both by SopB and by DNA (Bouet et al. 2007; Lim et al. 2005).

The δ protein of plasmid pSM19035 of *Streptococcus pyogenes* associates with the nucleoid. This pattern is altered in the presence of the ω CBF and centromere DNA, under which circumstances δ oscillates between the nucleoid and the cell poles, forming spiral-like structures. Accordingly, the protein polymerizes in vitro in response to ATP binding, and this polymerization is fully dependent on the presence of both DNA bearing the cognate centromere and the ω protein with which δ interacts (Pratto et al. 2008). Thus, the behaviour of δ may resemble that of Soj, a ParA homologue encoded by the *Bacillus subtilis* chromosome, that also polymerizes on DNA (Leonard et al. 2005). Soj binds to DNA non-specifically in vitro, and also associates with the nucleoid in vivo (Leonard et al. 2005; Marston and Errington 1999; Quisel et al. 1999). However, whether non-specific DNA binding by ParA proteins is universal is unclear, and the role that this non-specific binding might fulfil during DNA segregation remains uncertain (Castaing et al. 2008). ATP hydrolysis induces depolymerization of δ filaments indicating that the polymerization–depolymerization equilibrium is modulated by the nucleotide bound state of the protein. The dimeric crystal structure of δ suggests that ATP:ADP exchange can occur within the dimer without requiring dissociation into monomers (Pratto et al. 2008). By contrast, polymers of ParF and closely-related homologues (Fig. 4.2) are less prone to depolymerization in vitro and do not require DNA for filamentation (Barillà et al. 2005; Machón et al. 2007) suggesting that subfamilies of ParA proteins may possess discrete polymerization characteristics.

The ATPase activity of ParA proteins is stimulated by the cognate CBFs, as well as by non-specific DNA (Barillà et al. 2005; Davis et al. 1992; Fung et al. 2001; Libante et al. 2001; Pratto et al. 2008; Watanabe et al. 1992). A likely mechanism for this enhancement has emerged from studies of the N-terminal flexible domain of ParG and its interaction with ParF. Distinctly from its role in stimulation of ParF polymerization, the ParG tail possesses an arginine finger-like motif that promotes the ATPase activity of ParF by ~30-fold (Barillà et al. 2007). The motif may be part of a semi-flexible loop that intercalates into the ParF nucleotide binding pocket, analogous to arginine fingers in proteins such as human Ras-GAPs (Ahmadian et al. 1997;

Nadanaciva et al. 1999). Arginine finger loops stabilize the transition state during nucleotide hydrolysis by their partner proteins (Bos et al. 2007), and the same may be the case with ParF-ParG. Moreover, activation of nucleotide hydrolysis via an arginine finger loop may be a universally conserved, regulatory mechanism of ParA family members and their partner proteins (Barillà et al. 2007). This contention is supported by the observation that the mobile N-terminal tail in the ω CBF also is necessary for stimulation of ATPase hydrolysis by the δ protein (Pratto et al. 2008). Further study of the interaction between ParA homologues and their corresponding CBFs will provide crucial insights into the assembly, turnover, and mode of action of ParA filaments.

How might the dynamic polymerization of ParA proteins mediate plasmid segregation? Two plausible mechanisms focus on the motor-like action of ParA within the segrosome complex tethered at the centromere site (Barillà et al. 2005). Plasmid pairing at the mid-cell mediated by the segrosome is considered an initial step in the segregation process (Edgar et al. 2001; Funnell 2005; Jensen et al. 1998; Pratto et al. 2008; Ringaard et al. 2007b). One possibility is that, as with type II partitioning complexes (see Section 4.5), ATP-induced elongation of ParA polymers from paired segrosomes separates and pushes plasmids towards opposite cell poles. Alternatively, the coupled segrosomes that pair plasmids prior to segregation may remain relatively static until retraction of ParA polymers pulls the plasmids in different directions (Fig. 4.5) (Barillà et al. 2005). Support for the latter mechanism comes from observations that ParA polymers specified by one of the two chromosomes of *Vibrio cholerae* (Fig. 4.2) appear to pull the attached chromosome towards the cell pole in a process that superficially may mimic chromosome movement during mitosis in eukaryotic cells (Fogel and Waldor 2006). In either pushing or pulling scenarios, the ParA polymerization-depolymerization cycle likely requires an optimal balance between ATP-induced and ADP-mediated inhibition of polymerization (Barillà et al. 2007). To this end, the intrinsic ATPase activity of ParA must be stimulated at an appropriate rate and timepoint in the cell cycle by the cognate CBF. Additionally, the direct modulation of ParA filamentation by the CBF likely plays a fundamental role in the polymerization-depolymerization events. Indeed, there may be intriguing parallels between the effects of the CBFs on ParA polymerization and the action of microtubule associated proteins, formins, and other auxiliary protein factors that influence eukaryotic cytoskeletal dynamics (Machón et al. 2007).

4.5 An Actin-Like Polymerizing Protein in Type II Segrosomes

Although the genetic organizations of type I and II partition cassettes are very similar, the latter is distinguished particularly by the presence of a gene for a motor-like protein, ParM, that is structurally and functionally related to eukaryotic actin (Fig. 4.1) (Bork et al. 1992). Like actin, ParM is an ATPase (Jensen and Gerdes 1997). Moreover, ATP induces ParM polymerization in vitro (Møller-Jensen et al. 2002),

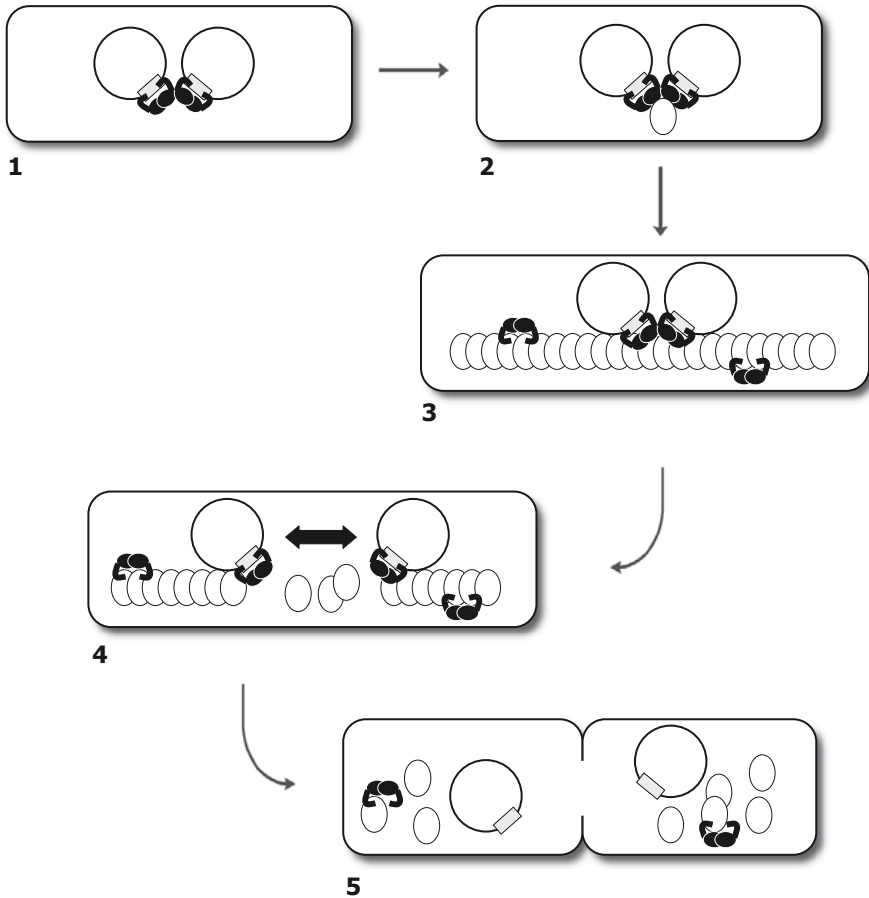


Fig. 4.5 Model for plasmid segregation mediated by ParA polymer retraction. (**Step 1**) The CBF (filled ovals) binds to the centromere (shaded box) generating a pre-segrosome with specific topology and causing plasmid pairing. (**Step 2**) apo-ParA (open oval) enters the mature segrosome by interaction with the CBF. (**Step 3**) ATP-induced polymerization of ParA causes bipolar filamentation of the protein. This filamentation may be modulated directly by the CBF. Stimulation of the ATPase activity of ParA by the CBF using an arginine finger motif provokes the conversion of ParA-ATP to an ADP-bound form which inhibits further polymerization. (**Step 4**) Depolymerization of the ParA polymers draws the plasmid pairs apart in opposite poleward directions. (**Step 5**) Septal closure traps the segregated plasmids in the daughter cells

although polymerization is even more proficient with GTP (Popp et al. 2008). ParM is also a more effective GTPase than ATPase (Popp et al. 2008). The crystal structures of ParM bound to ADP, GDP or non-hydrolyzable GMPPNP are very similar (Popp et al. 2008; van den Ent et al. 2002). Considering the properties of ParM-GTP mentioned above, this has led to a suggestion that GTP is the true functional ligand for ParM activity (Popp et al. 2008).

ParM filaments associated with plasmid DNA in vivo are observable in ~40% of fixed cells, with independent short clusters or dispersed filaments evident in the rest of the population (Moller-Jensen et al. 2002). Time-lapse microscopy of living cells revealed diffusive motions of plasmids that are increased by the presence of the type II partition operon, most likely encouraged by growth and retraction of the dynamic ParM filaments within the segrosomes (Campbell and Mullins 2007). For comparison, plasmids bearing type I cassettes have been reported to have more restricted intracellular plasmid movement, principally localizing to the mid-cell and quarter-cell positions under low copy number conditions (Derman et al. 2008; Li and Austin 2002; Li et al. 2004). ParM filaments grow bidirectionally, pushing plasmids attached to the polymer tips in opposite, pole-ward directions. The filaments disintegrate at the cell poles, allowing the plasmids to resume diffusive motions (Campbell and Mullins 2007). Release of nucleotide from the ends of the ParM polymers may induce a conformational change in ParM subunits at the polymer tips causing the dissociation of end subunits (Popp et al. 2008). The plasmids may subsequently be re-captured by ParM filaments. Thus, multiple successive rounds of ParM filamentation, plasmid movement, and polymer disassembly can occur in a single cell cycle suggesting that segregation mediated by type II complexes is uncoupled from the cell cycle, and may operate independently of host factors (Campbell and Mullins 2007). Electron cryomicroscopy of cells overproducing ParM revealed that the protein assembles into extensive arrays of closely-packed filament bundles. Under conditions that may more accurately mimic expression levels in the native partition operon, bundles of 3-to-5 ParM polymers located at the periphery of the nucleotide were detected (Salje et al. 2009).

Total internal reflection fluorescence (TIRF) microscopy demonstrated that ParM filaments polymerize bidirectionally in vitro when bound to ATP and are dynamically unstable, alternating between periods of regular growth and rapid shortening (Garner et al. 2004; Popp et al. 2007). The filaments are also dynamically unstable in vivo (Campbell and Mullins 2007). Dynamic instability may allow ParM filaments that are not inserted into the segrosome to disassemble, replenishing the pool of free ParM subunits for incorporation into productive bipolar filaments that are stabilized at both ends by caps of the ParR-*parC* complex. ParM polymers likely grow by an insertional polymerization mechanism in which new ParM subunits are inserted at the filament/ParR-*parC* interface, rather than within the polymers themselves (Garner et al. 2007; Moller-Jensen et al. 2003). Immobilization of the *parC* centromere on polystyrene beads, followed by addition of the ParR CBF, ParM and ATP, allowed reconstitution of the active segrosome in vitro (Garner et al. 2007). Under these conditions ParM formed dynamic radial asters up to ~3 μm in length around the bead. ParM filaments were much more elongated in the presence of the non-hydrolyzable ATP analogue, AMP-PNP, reflecting the repetitive growth and retraction of filaments that occurs in the presence of hydrolyzable nucleotide. Continuous ParM filaments also connected pairs of *parC*-coated beads. As the filaments elongated, these beads were propelled apart bidirectionally over considerable distances. By contrast, a ParM filament bound at one centromere may elongate and seek out another plasmid-bound ParM polymer

with which it then coalesces and can achieve bidirectional plasmid segregation (Garner et al. 2007). Thus, observations from in vivo and in vitro studies of ParM behaviour and dynamics correlate remarkably well.

Structural studies of ParM revealed that it forms double helical protofilaments with a longitudinal subunit spacing very similar to filamentous actin (F-actin) (Orlova et al. 2007; van den Ent et al. 2002). However, ParM filaments show left-handed twist compared to the right-handed twist of F-actin (Orlova et al. 2007; Popp et al. 2008). Moreover, domain-domain rotations within ParM subunits give rise to greater variability in twist angles than evident in F-actin. Furthermore, the subunit rotations in ParM polymers and in F-actin differ by $\sim 57^\circ$ generating entirely different subunit-subunit interfaces in the two filaments (Orlova et al. 2007). ParM and actin filaments are also distinguishable in vitro by the fast nucleation of ParM polymers, compared to more gradual actin nucleation (Garner et al. 2004). The elongation of actin filaments is unidirectional compared to the bidirectional growth of ParM polymers. Furthermore, instead of the treadmilling behaviour that is characteristic of F-actin, ParM shows dynamic instability. These features result in ParM filaments that are much shorter and with a more brief half-life than F-actin (Garner et al. 2004). Even more fundamentally, whereas ParM is implicated in DNA trafficking, actin cytoskeleton structures provide mechanical support to eukaryotic cells and are necessary for cell motility, among other functions, but are not directly involved in chromosome segregation. The evolutionary pathway that lead to prokaryotic and eukaryotic actin homologues with diverse functions continues to be explored (Becker et al. 2006; Lowe and Amos 2009).

4.6 Two Novel Classes of Segregation Complex: Tubulin and Coiled-Coil Partition Factors

Type I and II operons have been studied most extensively, but two other distinct partition cassettes have been recently identified. Type III segregation modules are located on large plasmids of the *Bacillus cereus* group of bacteria (Anand et al. 2008; Larsen et al. 2007). The cassettes comprise two genes, one of which encodes a distant homologue of eukaryotic tubulin, known either as TubZ or RepX. The preceding gene specifies the RepR product (Fig. 4.1). The TubZ/RepX proteins polymerize in vivo and the putative GTPase active site is necessary for normal polymerization (Akhtar et al. 2009; Larsen et al. 2007). Like tubulin, the TubZ polymers that assemble in vivo demonstrate treadmilling behaviour, elongating at one end while retracting at the opposite end (Larsen et al. 2007). The homologous TubZ and RepX proteins also polymerize in vitro into short, twisted two-stranded filaments in response to GTP binding. The proteins hydrolyze GTP strongly, as well as exhibiting weaker ATPase activity (Anand et al. 2008; Chen and Erickson 2008). As a result, TubZ/RepX exists predominantly in a GDP-bound form within filaments (Chen and Erickson 2008). Thus, like microtubules in which tubulin is also mainly present in the GDP-bound form, TubZ/RepX filaments exhibit dynamic instability,

depolymerizing spontaneously and rapidly *in vitro*. It has been proposed that the termini of TubZ/RepX polymers possess GTP caps that stabilize the GDP-bound species within the filaments which are otherwise prone to disassembly (Chen and Erickson 2008), akin to tubulin dynamics (Nogales and Wang 2006). The RepR protein negatively regulates intracellular TubZ levels (Larsen et al. 2007). Although the molecular basis for this regulation has yet to be clarified, RepR is a putative DNA binding protein and may autoregulate the *tubR-tubZ* cassette at the transcriptional level. RepX also binds DNA, albeit weakly and non-specifically (Anand et al. 2008). It is unclear whether this activity reflects the protein's role in plasmid segregation or replication: RepX is a probable replication initiation factor, as well as a plasmid partitioning protein (Tinsley and Khan 2006). The mechanism by which type III complexes segregate DNA remains to be determined, but it seems likely that the polymerizing activity of TubZ/RepX plays a crucial role in plasmid movement.

Type IV partition modules are exemplified by the *par* gene of staphylococcal plasmid pSK1. Homologous genes are located on plasmids in a diversity of Gram-positive species (Simpson et al. 2003). The *par* gene is unusual as, by contrast with type I–III systems, it appears to be the only gene necessary for plasmid stabilization (Fig. 4.1). The Par protein is predicted to bind DNA via an N-terminal helix-turn-helix motif, with a possible binding site(s) in the region 5' of *par* that is rich in repeat motifs and may act as a regulatory and/or centromeric sequence. A region of coiled-coil structure is predicted in the C-terminal region of the protein (Simpson et al. 2003). Further work is required to dissect the mode of plasmid stability mediated by type IV loci.

4.7 Perspectives

The plasmid segrosome provides a highly tractable framework to decipher the basis of bacterial DNA segregation. Although plasmid partition cassettes most commonly are of type I, the molecular mechanisms that drive the segregation of plasmids bearing these modules still remain to be fully solved. Nevertheless, it has been recently established that ParA homologues polymerize in response to nucleotide binding *in vitro*, assemble into filamentous cytoskeletal structures *in vivo*, and traffic plasmids to specific subcellular locations within the cell. What is the purpose of ParA polymerization during segregation? Like the actin-type ParM protein of type II cassettes, bipolar polymerization of motor-like ParA proteins may propel plasmids attached to the fibre tips in opposite directions. Alternatively, bidirectional disassembly of elongated ParA polymers may pull plasmids in opposite poleward directions.

There is now convincing evidence that CBFs encoded by type I cassettes load on to their centromeres and target plasmids to the mid-cell position where plasmid pairing occurs mediated by these pre-segrosomal structures. Is there an intracellular landmark to which the pre-segrosome is tethered at the cell centre? Data are also accumulating that pre-segrosomes are nucleoprotein superstructures in which the

CBF and centromeric DNA are arranged into intricate configurations. Considering the diversity in plasmid centromere organization (Fig. 4.3) and in the sequences of the CBFs that recognize them, it appears that pre-segrosomes assembled on different plasmids may possess substantially different architectures. The interface of ParA proteins with their cognate pre-segrosomes also may be specific. However, it is plausible that the subsequent actions of homologous ParA factors during segregation are broadly similar.

In the case of type II partition cassettes, the ParR CBF apparently does not modulate filamentation of the actin-like ParM protein. By contrast, CBFs encoded by type I modules not alone load on their centromeres, but also promote ParA protein polymerization. It is tempting to speculate that CBFs encoded by type I partition cassettes are related functionally to the plethora of protein factors that regulate eukaryote cytoskeletal dynamics. Moreover, the dual stimulation of ParA polymerization and of ParA ATPase activity by these CBFs must be integrated during plasmid segregation to ensure correct progression of the ParA polymerization-depolymerization cycle. Furthermore, whereas the type II segrosome apparently can assemble and disassemble repeatedly during the bacterial cell cycle, it is unknown whether the same pertains for type I complexes, or whether a single partitioning event is necessary and sufficient for accurate plasmid maintenance. Further characterization of prototypical segregation complexes, as well as of informative variant systems, will continue to provide key mechanistic insights into genome segregation, a basic cellular process.

Acknowledgements We apologize to authors whose valuable work could not be discussed in detail or cited due to space limitations. Work in the authors' laboratories is supported by the Biotechnology and Biological Sciences Research Council (grants BB/G003114/1 and BB/F012004/1 to FH and DB, respectively); the Medical Research Council (grant G0500588 to FH and DB, and grant G0801162 to DB); and, by European Union contract LSHM-CT-2005-019023 to FH.

References

- Adachi S, Hori K, Hiraga S (2006) Subcellular positioning of F plasmid mediated by dynamic localization of SopA and SopB. *J Mol Biol* 356:850–863
- Ahmadian MR, Stege P, Scheffzek K, Wittinghofer A (1997) Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. *Nat Struct Biol* 4:686–689
- Akhtar P, Anand SP, Watkins SC, Khan SA (2009) The tubulin-like RepX protein encoded by the pXO1 plasmid forms polymers in vivo in *Bacillus anthracis*. *J Bacteriol* 191:2493–2500
- Anand SP, Akhtar P, Tinsley E, Watkins SC, Khan SA (2008) GTP-dependent polymerization of the tubulin-like RepX replication protein encoded by the pXO1 plasmid of *Bacillus anthracis*. *Mol Microbiol* 67:881–890
- Barillà D, Hayes F (2003) Architecture of the ParF-ParG protein complex involved in procaryotic DNA segregation. *Mol Microbiol* 49:487–499
- Barillà D, Rosenberg MF, Nobbmann U, Hayes F (2005) Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *EMBO J* 24:1453–1464
- Barillà D, Carmelo E, Hayes F (2007) The tail of the ParG DNA segregation protein remodels ParF polymers and enhances ATP hydrolysis via an arginine finger-like motif. *Proc Natl Acad Sci USA* 104:1811–1816

- Bartosik AA, Jagura-Burdzy G (2005) Bacterial chromosome segregation. *Acta Biochim Pol* 52:1–34
- Becker E, Herrera NC, Gunderson FQ, Derman AI, Dance AL, Sims J, Larsen RA, Pogliano J (2006) DNA segregation by the bacterial actin AlfA during *Bacillus subtilis* growth and development. *EMBO J* 25:5919–5931
- Bork P, Sander C, Valencia A (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc Natl Acad Sci USA* 89:7290–7294
- Bos JL, Rehmann H, Wittinghofer A (2007) GEFs and GAPs: critical elements in the control of small G proteins. *Cell* 129:865–877
- Bouet JY, Funnell BE (1999) P1 ParA interacts with the P1 partition complex at *parS* and an ATP-ADP switch controls ParA activities. *EMBO J* 18:1415–1424
- Bouet JY, Ah-Seng Y, Benmeradi N, Lane D (2007) Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Mol Microbiol* 63:468–481
- Breüner A, Jensen RB, Dam M, Pedersen S, Gerdes K (1996) The centromere-like *parC* locus of plasmid R1. *Mol Microbiol* 20:581–592
- Campbell CS, Mullins RD (2007) In vivo visualization of type II plasmid segregation: bacterial actin filaments pushing plasmids. *J Cell Biol* 179:1059–1066
- Carmelo E, Barillà D, Golovanov AP, Lian LY, Derome A, Hayes F (2005) The unstructured N-terminal tail of ParG modulates assembly of a quaternary nucleoprotein complex in transcription repression. *J Biol Chem* 280:28683–28691
- Castaing JP, Bouet JY, Lane D (2008) F plasmid partition depends on interaction of SopA with non-specific DNA. *Mol Microbiol* 70:1000–1011
- Chattoraj DK (2000) Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol Microbiol* 37:467–476
- Chen Y, Erickson HP (2008) In vitro assembly studies of FtsZ/tubulin-like proteins (TubZ) from *Bacillus* plasmids: evidence for a capping mechanism. *J Biol Chem* 283:8102–8109
- Cooper JA, Sept D (2008) New insights into mechanism and regulation of actin capping protein. *Int Rev Cell Mol Biol* 267:183–206
- Dabrazhynetskaya A, Sergueev K, Austin S (2005) Species and incompatibility determination within the P1*par* family of plasmid partition elements. *J Bacteriol* 187:5977–5983
- Dabrazhynetskaya A, Brendler T, Ji X, Austin S (2009) Switching protein-DNA recognition specificity by single amino acid substitutions in the P1 *par* family of plasmid partition elements. *J Bacteriol* 191:1126–1131
- Dam M, Gerdes K (1994) Partitioning of plasmid R1. Ten direct repeats flanking the *parA* promoter constitute a centromere-like partition site *parC*, that expresses incompatibility. *J Mol Biol* 236:1289–1298
- Davey MJ, Funnell BE (1994) The P1 plasmid partition protein ParA. A role for ATP in site-specific DNA binding. *J Biol Chem* 269:29908–29913
- Davis MA, Martin KA, Austin SJ (1992) Biochemical activities of the ParA partition protein of the P1 plasmid. *Mol Microbiol* 6:1141–1147
- Davis MA, Radnedge L, Martin KA, Hayes F, Youngren B, Austin SJ (1996) The P1 ParA protein and its ATPase activity play a direct role in the segregation of plasmid copies to daughter cells. *Mol Microbiol* 21:1029–1036
- de la Hoz AB, Ayora S, Sitkiewicz I, Fernández S, Pankiewicz R, Alonso JC, Ceglowski P (2000) Plasmid copy-number control and better-than-random segregation genes of pSM19035 share a common regulator. *Proc Natl Acad Sci USA* 97:728–733
- de la Hoz AB, Pratto F, Misselwitz R, Speck C, Weihofen W, Welfle K, Saenger W, Welfle H, Alonso JC (2004) Recognition of DNA by ω protein from the broad-host range *Streptococcus pyogenes* plasmid pSM19035: analysis of binding to operator DNA with one to four heptad repeats. *Nucleic Acids Res* 32:3136–3147
- del Solar G, Espinosa M (2000) Plasmid copy number control: an ever-growing story. *Mol Microbiol* 37:492–500
- Delbrück H, Ziegelin G, Lanka E, Heinemann U (2002) An Src homology 3-like domain is responsible for dimerization of the repressor protein Korb encoded by the promiscuous IncP plasmid RP4. *J Biol Chem* 277:4191–4198

- Derman AI, Lim-Fong G, Pogliano J (2008) Intracellular mobility of plasmid DNA is limited by the ParA family of partitioning systems. *Mol Microbiol* 67:935–946
- Derome A, Hoischen C, Bussiek M, Grady R, Adamczyk M, Kędzierska B, Diekmann S, Barillà D, Hayes F (2008) Centromere anatomy in the multidrug-resistant pathogen *Enterococcus faecium*. *Proc Natl Acad Sci USA* 105:2151–2156
- Dmowski M, Sitkiewicz I, Ceglowski P (2006) Characterization of a novel partition system encoded by the δ and ω genes from the streptococcal plasmid pSM19035. *J Bacteriol* 188:4362–4372
- Ebersbach G, Gerdes K (2001) The double *par* locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. *Proc Natl Acad Sci USA* 98:15078–15083
- Ebersbach G, Gerdes K (2005a) Plasmid segregation mechanisms. *Annu Rev Genet* 39:453–479
- Ebersbach G, Gerdes K (2005b) Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Mol Microbiol* 52:385–398
- Ebersbach G, Ringgaard S, Moller-Jensen J, Wang Q, Sherratt DJ, Gerdes K (2006) Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. *Mol Microbiol* 61:1428–1442
- Edgar R, Chatteraj DK, Yarmolinsky M (2001) Pairing of P1 plasmid partition sites by ParB. *Mol Microbiol* 42:1363–1370
- Fogel MA, Waldor MK (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev* 20:3269–3282
- Fothergill TJG, Barillà D, Hayes F (2005) Protein diversity confers specificity in plasmid segregation. *J Bacteriol* 187:2651–2661
- Friedman SA, Austin SJ (1988) The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon. *Plasmid* 19:103–112
- Fung E, Bouet JY, Funnell BE (2001) Probing the ATP-binding site of P1 ParA: partition and repression have different requirements for ATP binding and hydrolysis. *EMBO J* 20:4901–4911
- Funnell BE (1988) Participation of *Escherichia coli* integration host factor in the P1 plasmid partition system. *Proc Natl Acad Sci USA* 85:6657–6661
- Funnell BE (1991) The P1 plasmid partition complex at *parS*. The influence of *Escherichia coli* integration host factor and of substrate topology. *J Biol Chem* 266:14328–14337
- Funnell BE (2005) Partition-mediated plasmid pairing. *Plasmid* 53:119–125
- Garner EC, Campbell CS, Mullins RD (2004) Dynamic instability in a DNA-segregating prokaryotic actin homolog. *Science* 306:1021–1025
- Garner EC, Campbell CS, Weibel DB, Mullins RD (2007) Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. *Science* 315:1270–1274
- Golovanov AP, Barillà D, Golovanova M, Hayes F, Lian LY (2003) ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Mol Microbiol* 50:1141–1153
- Hatano T, Yamaichi Y, Niki H (2007) Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Mol Microbiol* 64:1198–1213
- Hayes F (2000) The partition system of multidrug resistance plasmid TP228 includes a novel protein that epitomizes an evolutionarily-distinct subgroup of the ParA superfamily. *Mol Microbiol* 37:528–541
- Hayes F (2003a) The function and organization of plasmids. *Methods Mol Biol* 235:1–17
- Hayes F (2003b) Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science* 301:1496–1499
- Hayes F, Austin SJ (1993) Specificity determinants of the P1 and P7 plasmid centromere analogs. *Proc Natl Acad Sci USA* 90:9228–9232
- Hayes F, Austin S (1994) Topological scanning of the P1 plasmid partition site. *J Mol Biol* 243:190–198
- Hayes F, Barillà D (2006a) The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat Rev Microbiol* 4:133–143
- Hayes F, Barillà D (2006b) Assembling the bacterial segrosome. *Trends Biochem Sci* 31:247–250
- Hayes F, Davis MA, Austin SJ (1993) Fine-structure analysis of the P7 plasmid partition site. *J Bacteriol* 175:3443–3451
- Hayes F, Radnedge L, Davis MA, Austin SJ (1994) The homologous operons for P1 and P7 plasmid partition are autoregulated from dissimilar operator sites. *Mol Microbiol* 11:249–260

- Hoischen C, Bolshoy A, Gerdes K, Diekmann S (2004) Centromere *parC* of plasmid R1 is curved. *Nucleic Acids Res* 32:5907–5915
- Hoischen C, Bussiek M, Langowski J, Diekmann S (2008) *Escherichia coli* low-copy-number plasmid R1 centromere *parC* forms a U-shaped complex with its binding protein ParR. *Nucleic Acids Res* 36:607–615
- Jensen RB, Gerdes K (1997) Partitioning of plasmid R1. The ParM protein exhibits ATPase activity and interacts with the centromere-like ParR-*parC* complex. *J Mol Biol* 269:505–513
- Jensen RB, Dam M, Gerdes K (1994) Partitioning of plasmid R1. The *parA* operon is autoregulated by ParR and its transcription is highly stimulated by a downstream activating element. *J Mol Biol* 236:1299–1309
- Jensen RB, Lurz R, Gerdes K (1998) Mechanism of DNA segregation in prokaryotes: replicon pairing by *parC* of plasmid R1. *Proc Natl Acad Sci USA* 95:8550–8555
- Khare D, Ziegelin G, Lanka E, Heinemann U (2004) Sequence-specific DNA binding determined by contacts outside the helix-turn-helix motif of the ParB homolog KorB. *Nat Struct Mol Biol* 11:656–663
- Koonin EV (1993) A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J Mol Biol* 229:1165–1174
- Larsen RA, Cusumano C, Fujioka A, Lim-Fong G, Patterson P, Pogliano J (2007) Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in *Bacillus thuringiensis*. *Genes Dev* 21:1340–1352
- Leonard TA, Butler PJ, Lowe J (2005) Bacterial chromosome segregation: structure and DNA binding of the Soj dimer – a conserved biological switch. *EMBO J* 24:270–282
- Li Y, Austin S (2002) The P1 plasmid is segregated to daughter cells by a ‘capture and ejection’ mechanism coordinated with *Escherichia coli* cell division. *Mol Microbiol* 46:63–74
- Li Y, Dabrazhynetskaya A, Youngren B, Austin S (2004) The role of Par proteins in the active segregation of the P1 plasmid. *Mol Microbiol* 53:93–102
- Libante V, Thion L, Lane D (2001) Role of the ATP-binding site of SopA protein in partition of the F plasmid. *J Mol Biol* 314:387–399
- Lim GE, Derman AI, Pogliano J (2005) Bacterial DNA segregation by dynamic SopA polymers. *Proc Natl Acad Sci USA* 102:17658–17663
- Lowe J, Amos LA (2009) Evolution of cytomotive filaments: the cytoskeleton from prokaryotes to eukaryotes. *Int J Biochem Cell Biol* 41:323–329
- Lynch AS, Wang JC (1995) SopB protein-mediated silencing of genes linked to the *sopC* locus of *Escherichia coli* F plasmid. *Proc Natl Acad Sci USA* 92:1896–1900
- Machón C, Fothergill TJG, Barillà D, Hayes F (2007) Promiscuous stimulation of ParF protein polymerization by heterogeneous centromere binding factors. *J Mol Biol* 374:1–8
- Marston AL, Errington J (1999) Dynamic movement of the ParA-like Soj protein of *B. subtilis* and its dual role in nucleoid organization and developmental regulation. *Mol Cell* 4:673–682
- Moller-Jensen J, Jensen RB, Lowe J, Gerdes K (2002) Prokaryotic DNA segregation by an actin-like filament. *EMBO J* 21:3119–3127
- Moller-Jensen J, Borch J, Dam M, Jensen RB, Roepstorff P, Gerdes K (2003) Bacterial mitosis: ParM of plasmid R1 moves plasmid DNA by an actin-like insertional polymerization mechanism. *Mol Cell* 12:1477–1487
- Moller-Jensen J, Ringgard S, Mercogliano CP, Gerdes K, Lowe J (2007) Structural analysis of the ParR/*parC* plasmid partition complex. *EMBO J* 26:4413–4422
- Motallebi-Veshareh M, Rouch DA, Thomas CM (1990) A family of ATPases involved in active partitioning of diverse bacterial plasmids. *Mol Microbiol* 4:1455–1463
- Murayama K, Orth P, de la Hoz AB, Alonso JC, Saenger W (2001) Crystal structure of ω transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 Å resolution. *J Mol Biol* 314:789–796
- Nadanaciva S, Weber J, Wilke-Mounts S, Senior AE (1999) Importance of F₁-ATPase residue α -Arg-376 for catalytic transition state stabilization. *Biochemistry* 38:15493–15499
- Nogales E, Wang HW (2006) Structural mechanisms underlying nucleotide-dependent self-assembly of tubulin and its relatives. *Curr Opin Struct Biol* 16:221–229

- Nordstrom K (2006) Plasmid R1 – replication and its control. *Plasmid* 55:1–26
- Orlova A, Garner EC, Galkin VE, Heuser J, Mullins RD, Egelman EH (2007) The structure of bacterial ParM filaments. *Nat Struct Mol Biol* 14:921–926
- Pogliano J (2008) The bacterial cytoskeleton. *Curr Opin Cell Biol* 20:19–27
- Popp D, Yamamoto A, Iwasa M, Narita A, Maeda K, Maéda Y (2007) Concerning the dynamic instability of actin homolog ParM. *Biochem Biophys Res Commun* 353:109–114
- Popp D, Narita A, Oda T, Fujisawa T, Matsuo H, Nitanaï Y, Iwasa M, Maeda K, Onishi H, Maéda Y (2008) Molecular structure of the ParM polymer and the mechanism leading to its nucleotide-driven dynamic instability. *EMBO J* 27:570–579
- Pratto F, Cicek A, Weihofen WA, Lurz R, Saenger W, Alonso JC (2008) *Streptococcus pyogenes* pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on *parS* DNA during plasmid segregation. *Nucleic Acids Res* 36:3676–3689
- Quisel JD, Lin DC, Grossman AD (1999) Control of development by altered localization of a transcription factor in *B. subtilis*. *Mol Cell* 4:665–672
- Radnedge L, Davis MA, Austin SJ (1996) P1 and P7 plasmid partition: ParB protein bound to its partition site makes a separate discriminator contact with the DNA that determines species specificity. *EMBO J* 15:1155–1162
- Radnedge L, Youngren B, Davis M, Austin S (1998) Probing the structure of complex macromolecular interactions by homolog specificity scanning: the P1 and P7 plasmid partition systems. *EMBO J* 17:6076–6085
- Ringaard S, Ebersbach G, Borch J, Gerdes K (2007a) Regulatory cross-talk in the double *par* locus of plasmid pB171. *J Biol Chem* 282:3134–3145
- Ringaard S, Lowe J, Gerdes K (2007b) Centromere pairing by a plasmid-encoded type I ParB protein. *J Biol Chem* 282:28216–28225
- Rodionov O, Yarmolinsky M (2004) Plasmid partitioning and the spreading of P1 partition protein ParB. *Mol Microbiol* 52:1215–1223
- Rodionov O, Lobočka M, Yarmolinsky M (1999) Silencing of genes flanking the P1 plasmid centromere. *Science* 283:546–549
- Rothfield L, Taghbalout A, Shih YL (2005) Spatial control of bacterial division-site placement. *Nat Rev Microbiol* 3:959–968
- Salje J, Lowe J (2008) Bacterial actin: architecture of the ParMRC plasmid DNA partitioning complex. *EMBO J* 27:2230–2238
- Salje J, Zuber B, Lowe J (2009) Electron cryomicroscopy of *E. coli* reveals filament bundles involved in plasmid DNA segregation. *Science* 323:509–512
- Schumacher MA (2007) Structural biology of plasmid segregation proteins. *Curr Opin Struct Biol* 17:103–109
- Schumacher MA (2008) Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. *Biochem J* 412:1–18
- Schumacher MA, Funnell BE (2005) Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature* 438:516–519
- Schumacher MA, Glover TC, Brzoska AJ, Jensen SO, Dunham TD, Skurray RA, Firth N (2007a) Segrosome structure revealed by a complex of ParR with centromere DNA. *Nature* 450:1268–1271
- Schumacher MA, Mansoor A, Funnell BE (2007b) Structure of a four-way bridged ParB-DNA complex provides insight into P1 segrosome assembly. *J Biol Chem* 282:10456–10464
- Sergueev K, Dabrazhynetskaya A, Austin S (2005) Plasmid partition system of the P1*par* family from the pWR100 virulence plasmid of *Shigella flexneri*. *J Bacteriol* 187:3369–3373
- Simpson AE, Skurray RA, Firth N (2003) A single gene on the staphylococcal multiresistance plasmid pSK1 encodes a novel partitioning system. *J Bacteriol* 185:2143–2152
- Summers D (1998) Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol Microbiol* 29:1137–1145
- Thanbichler M, Shapiro L (2006) Chromosome organization and segregation in bacteria. *J Struct Biol* 156:292–303
- Thomas CM (ed) (2000) The horizontal gene pool. Harwood Academic, Amsterdam, The Netherlands

- Tinsley E, Khan SA (2006) A novel FtsZ-like protein is involved in replication of the anthrax toxin-encoding pXO1 plasmid in *Bacillus anthracis*. *J Bacteriol* 188:2829–2835
- van den Ent F, Moller-Jensen J, Amos LA, Gerdes K, Lowe J (2002) F-actin-like filaments formed by plasmid segregation protein ParM. *EMBO J* 21:6935–6943
- Vecchiarelli AG, Schumacher MA, Funnell BE (2007) P1 partition complex assembly involves several modes of protein-DNA recognition. *J Biol Chem* 282:10944–10952
- Watanabe E, Wachi M, Yamasaki M, Nagai K (1992) ATPase activity of SopA, a protein essential for active partitioning of F plasmid. *Mol Gen Genet* 234:346–352
- Weihs WA, Cicek A, Pratto F, Alonso JC, Saenger W (2006) Structures of ω repressors bound to direct and inverted DNA repeats explain modulation of transcription. *Nucleic Acids Res* 34:1450–1458
- Zampini M, Derome A, Bailey SES, Barillà D, Hayes F (2009) Recruitment of the ParG segregation protein to different affinity DNA sites. *J Bacteriol* 191:3832–3841

Chapter 5

Nucleoid Structure and Segregation

Conrad L. Woldringh

Abstract One of the fundamental differences between eukaryotic and bacterial cell cycles is the possibility of re-replication in bacteria. The organization of the bacterial DNA in the nucleoid and the mechanism of its segregation make it possible that during an ongoing round of replication, new initiations can take place. In contrast to the protein-rich, nucleosomal structure of eukaryotic DNA, the bacterial chromosome is thought to consist of thousands of supercoiled, branched segments that are compacted in a protein-poor nucleoid phase through non-specific volume-exclusion interactions with cytoplasmic proteins. Recent findings on the movement of fluorescent loci on both arms (replichores) of the *Escherichia coli* chromosome and on the dynamics of DNA polymerases have given a coherent picture of nucleoid organization. While the chromosome is replicated bidirectionally from a single origin by independent replisomes, the daughter strands separate and move past the bulk of unreplicated DNA in a simultaneous replication/segregation process. Repulsive, entropic forces probably promote the segregation of DNA daughter strands. In *E. coli* the two replichores of each separating daughter strand can move apart at different velocities and become positioned in two halves of the newly synthesized nucleoid with the origin in between. In contrast, in *Caulobacter crescentus* and *Vibrio cholerae* (*oriCI*), one origin is kept near the old cell pole and both replichores move together with the other origin at the tip of the newly synthesized nucleoid. More accurate determinations of the dynamics of DNA loci under various growth conditions and during growth inhibition (run-off DNA synthesis) may be expected to establish whether entropic forces are sufficient for DNA segregation or whether dedicated biological mechanisms have to give an extra drift, for instance in initial origin separation or in establishing the division of the nucleoid.

Keywords Bacterial chromosome • chromosome arms (replichores) • entropic repulsion of DNA strands • nucleoid segregation • phase separation • supercoiled polymer-network • volume-exclusion interactions

C.L. Woldringh (✉)

Molecular Cytology, Faculty of Science, Swammerdam Institute for Life Sciences,
University of Amsterdam, Amsterdam, The Netherlands
e-mail: c.l.woldringh@uva.nl

5.1 Introduction

In bacteria chromosomes usually contain a single origin of replication from which two replisomes start to replicate the chromosome arms in opposite directions (Fig. 5.1a). Fast growing *E. coli* cells (Fig. 5.1b) divide more quickly than they can produce a complete chromosome, as the replication time (C-period) may be longer than the generation time (T_d). To produce complete, viable cells, the cell has to initiate DNA replication in step with cell growth and at the same frequency as it divides, i.e. once per cell cycle (Helmstetter and Cooper 1968; Hansen et al. 1991). During fast growth this implies that initiation occurs on a chromosome that is still carrying out one or even two round(s) of replication. This so-called multi-fork replication during overlapping C-periods is only possible if newly replicated daughter strands are separated as soon as they are synthesized, in a concurrent replication and segregation process (right panels in Fig. 5.1a, b). This avoids the possibility of entanglement of newly replicated strands and avoids the necessity for a disentangling mechanism that has to recognize the hierarchy of the replicated strands (Donachie et al. 1995).

Eukaryotic cells have no mechanism for beginning a new DNA replication cycle before finishing the last one. Such re-replications or overlapping S-phases are not possible, but also not necessary because eukaryotes can start DNA replication at many origins. Chromosome replication thus occurs in a period (S + G₂/M) that is always shorter than the cell's generation time. In a first stage of segregation (Fig. 5.1c), sister chromatids are aligned, held together by cohesins (Nasmyth 2002) and subsequently disentangled by a mechanism that may involve chromosome condensation by condensins (Marko and Siggia 1997); this mechanism is still not well understood (Guacci 2007). In a second stage of segregation (Fig. 5.1d), microtubules attach to a single, kinetochore built on the aligned centromeres of the compact chromosomes. This mechanism, mitosis, is unsuitable for separating replicating chromosomes, let alone chromosomes that are in a state of more than one round of replication like in *E. coli* (Fig. 5.1b).

The use of fluorescence microscopy to study bacterial strains with fluorescent DNA markers has greatly advanced our understanding of bacterial chromosome organization (Gordon et al. 1997; Niki et al. 2000; Wang et al. 2005; Nielsen et al. 2006a). However, the finding of rapid movements of fluorescently labeled origins has led some authors to consider that bacterial chromosome segregation occurs by a mechanism similar to that in eukaryotes. This also led them to adopt a eukaryotic terminology for the bacterial cell cycle (e.g. Gitai et al. 2005; Fogel and Waldor 2006). This terminology, however, conceals a fundamental difference between the two systems of DNA replication: the impossibility in eukaryotic cells of concurrent DNA replication and segregation (i.e. overlapping S-phases) and the absence in bacteria of a G₂/M period in which chromosome condensation takes place. Thus, as has been suggested earlier (Nasmyth 2002; Woldringh and van Driel 1999), bacterial segregation can better be compared with the first stage of disentanglement of the eukaryotic chromatids (Fig. 5.1c), than with the second stage, mitosis

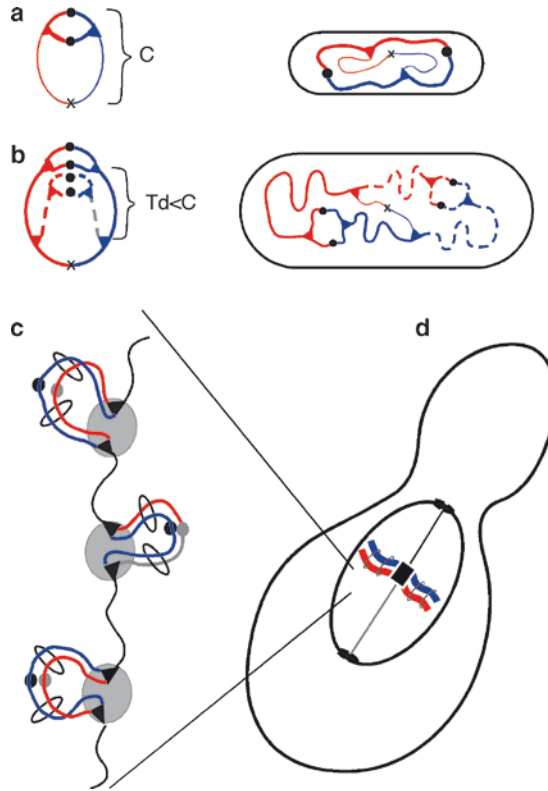


Fig. 5.1 Schematic representation of chromosome replication and segregation in pro- and eukaryotes. Cell sizes are not to scale. Origins of replication are indicated by circles, replisomes by triangles. (a) The circular chromosome of a bacterium like *E. coli* is replicated during the C-period (*slow growth conditions*). Different replichores are in black and grey; different daughter strands are in full or dashed lines. (b) During fast growth, re-replications from multiple origins occur every generation time, T_d , which is shorter than the C-period. Right panels in (a) and (b): simultaneous replication and segregation of the origins. Note larger cell in B as predicted by the Helmstetter-Cooper model (1968). Different daughter strands are in full and dashed lines. (c) The linear chromosome of a eukaryotic cell like yeast is replicated from many origins. Different daughter strands (chromatids) are in black and grey. The nucleosomal structure of the DNA is not indicated. The replisomes occur in replication factories (grey circles). Replication is coupled with cohesion by cohesin complexes indicated by ovals. Groups of replisome pairs can combine in a single factory (see Kitamura et al. 2006). (d) Sister chromatids are disentangled, aligned and condensed by condensin complexes in the yeast nucleus. Final segregation occurs after microtubules have attached to the kinetochore structure (black square) and cohesins have become proteolyzed

(Fig. 5.1d), in which the condensed chromatids are pulled to opposite cell poles with the help of microtubules.

In this review I will first discuss cytological and physical aspects of the structure of the bacterial nucleoid both in situ and in isolated nucleoids. In subsequent sections a rather coherent picture of segregation is given based on recent findings of the dynamics of fluorescent DNA markers and replisomes. I will argue that replication

occurs in the outer border of the nucleoid where entropic, repulsive interactions between newly replicated segments of the chromosome cause their segregation.

5.2 Nucleoid Structure

5.2.1 *Conclusions Based on Light Microscopy, Electron Microscopy and Cryoelectron Tomography*

That bacterial DNA is found in visible structures called nucleoids is clear for fast-growing, relatively large *E. coli* cells. This was demonstrated in 1956 in elegant phase-contrast images by Mason and Powelson (1956) and have to my knowledge only been reproduced with the same quality by Hironori Niki (National Institute of Genetics, Japan, unpublished). These timelapse sequences of cells embedded in gelatin indicate that the bacterial nucleoid is a low mass-density region that enlarges and duplicates in step with cell growth and division. In smaller cells, like slow-growing *E. coli* or *Caulobacter crescentus*, distinct nucleoids are difficult to see using either phase contrast or fluorescence microscopy. Indeed *Caulobacter* cells stained with DAPI (6-diamidino-2-phenylindole dihydrochloride hydrate) were first reported to have no chromosome-free regions and thus no nucleoids (Jensen and Shapiro 1999; Jensen 2006). However, combined phase-contrast and fluorescence microscopy suggests that there is a cytoplasmic phase along the cell border and that the DNA is not fully dispersed throughout the cell (Fig. 5.2a, left panel). The phase separation in the small cell is made more obvious by fixation with osmium tetroxide (Fig. 5.2a, right panel) and by electron microscopy of thin sections of both *C. crescentus* (Poindexter 1964) and *E. coli* cells (Woldringh et al. 1977). Such electron micrographs show occasional fibrils which penetrate from the nucleoid into the cytoplasm, possibly connecting even to the plasma membrane (van Iterson 1965). However, the phase separation in osmium-tetroxide fixed cells could be an artifact because fixation with glutaraldehyde does not reveal such a clear separation between a fibrillar nucleoid and a granular cytoplasm (see for reviews Woldringh and Nanninga 1985; Robinow and Kellenberger 1994).

Does this mean that in small, live cells the DNA is dispersed throughout the cytoplasm and that there is no separate cytoplasmic phase as seen in the large, live cells of Mason and Powelson (1956)? It was hoped that this question could be answered by the introduction of cryofixation techniques. However, in freeze-fracture preparations nucleoids were not always visible (Nanninga 1969). Also with the more recently developed technique of cryoelectron tomography a nucleoid is not always present, e.g. in tomograms of the very small bacterium *Spiroplasma melliferum* (Fig. 5.2e; Ortiz et al. 2006). It is therefore relevant that ribosome-free regions with a different texture could be seen in vitrified sections of *Deinococcus radiodurans* (Eltsov and Dubochet 2005) and in cryoelectron tomograms of *Bdellovibrio bacteriovorus* (Fig. 5.2d; Borgnia et al. 2008). The comparison of cell and nucleoid sizes of these different bacteria (Fig. 5.2b–e) shows that nucleoid segregation in the

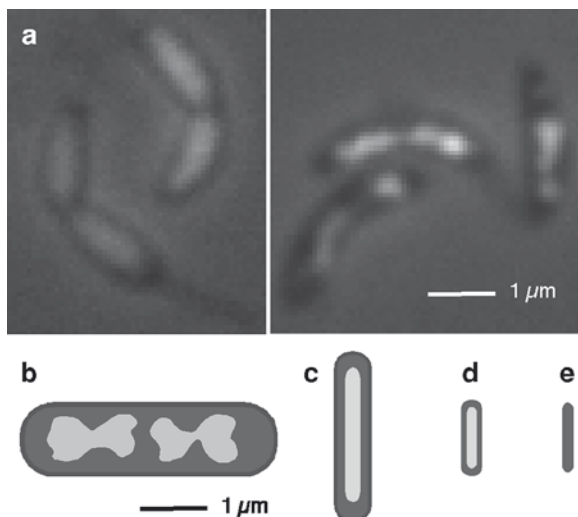


Fig. 5.2 Comparison of cell sizes and nucleoids in different bacteria. (a) *C. crescentus* CB15N grown in PYE medium at 30°C into stationary phase. Cells stained with DAPI (1 $\mu\text{g/mL}$) and photographed by combined phase-contrast and fluorescence illumination. Left panel: living cells; right panel: cells fixed with 0.1% osmium tetroxide. (b–d) Schematic drawing to scale of: (b) *E. coli* B/rH grown in LB-medium (doubling time of 20 min; cell volume 2.70 μm^3); (c) *E. coli* B/rH grown in alanine medium (doubling time of 150 min; cell volume 0.46 μm^3 ; DNA concentration 0.011 g/mL cell volume); (d) *Bdellovibrio bacteriovorus* (cell volume 0.08 μm^3 ; DNA concentration 0.043 g/mL cell volume); (e) *Spiroplasma melliferum* (cell volume 0.02 μm^3 ; DNA concentration 0.052 g/mL cell volume). See text for references

small cells of *Bdellovibrio* and *Spiroplasma* takes place over a distance equal to the diameter of a fast-growing *E. coli* cell. In the three cell types of Fig. 5.2c–e, the DNA concentration increases about five-fold. It remains a question whether the absence of a visible nucleoid in the small cell of *Spiroplasma* (Fig. 5.2e) can be ascribed to its high DNA concentration (see legend Fig. 5.2) or to an effect of the cryo-fixation procedure. For instance, it has been noted in the past that freezing techniques failed to preserve adhesion sites between inner and outer membranes of plasmolyzed *E. coli* cells (Bayer 1991). Considering the physical forces that are involved in DNA compaction as described in the next section, it could be envisaged that during the freezing process, which of necessity begins in the extracellular space, DNA can become dispersed.

5.2.2 Theoretical Prediction of a Phase Separation Between Nucleoid and Cytoplasm

The principle of phase separation between a protein-poor nucleoid and a protein-rich cytoplasm (Valkenburg and Woldringh 1984) is based on macromolecular crowding as first proposed by Zimmerman and Murphy (1996). Odijk (1998) predicted

this phenomenon on theoretical grounds based on polymer-physical considerations. Assuming that the bacterial chromosome occurs in the form of a huge branched, plectonemic supercoil in a cell containing a large number of soluble, negatively charged proteins (Fig. 5.3a), he derived two equations for the free energy of excluded volume interactions: (i) the free energy of self-interaction between the supercoiled DNA segments (also called Kuhn segments), and (ii) the free energy of cross interactions between proteins and DNA (also called depletion energy). The calculations predict a tenfold larger energy for the protein-DNA cross-interactions than for the DNA self-interactions. A thermodynamic equilibrium for this unstable situation is obtained by a phase separation between a more compact nucleoid in which fewer proteins interact with the DNA and a cytoplasmic phase in which the proteins merely interact with each other. This results in an equilibrium situation with a lower free energy than when the DNA would be dispersed throughout the cell (see for a discussion of these calculations Woldringh and Odijk 1999; Woldringh 2002). The very small cells of *S. melliferum* in which no phase separation was observed (see Fig. 5.2e) may not contain enough soluble proteins to establish such a phase separation.

This phase separation (Fig. 5.3) is important for understanding of nucleoid structure and segregation. It implies that the nucleoid volume is significantly smaller than cell volume, that there is a “surface” or border region between the two phases and that the nucleoid has different physical and chemical properties than when dispersed throughout the cell. It remains unknown what energy it costs for a loop of DNA to stretch out into the cytoplasm as depicted in Fig. 5.3c and as sometimes visible in electron micrographs of thin sections (van Iterson 1965). Therefore, the implications of the phase separation for properties at the interface between nucleoid and cytoplasm are uncertain. As will be discussed below, the process of replication may occur in this border region where the dynamics of DNA strands could be higher than inside the congested nucleoid (Fig. 5.3b). Also, the process of transcription (Fig. 5.3c) may be localized in the nucleoid periphery as reported by Jin and Cabrera (2006) and as has been proposed more than 30 years ago by Maaløe and Kjeldgaard (1966); see their Figure 7.1).

5.2.3 DNA Binding Proteins

According to the above physical considerations the collective action of cytoplasmic proteins represents a large and aspecific compaction force that can probably not be attained by the number and action of specific DNA-binding proteins like HU, IHF, H-NS and FIS (Azam et al. 1999; Zimmerman 2006a; see for reviews Stavans and Oppenheim 2006; Luijsterburg et al. 2006). However, these proteins that can bridge, bend and wrap DNA, may have large, local effects on DNA behavior. By binding to specific DNA segments they can change the physical properties of the DNA like persistence length and superhelicity, thereby changing the parameters of self-interactions between segments and cross-interactions with soluble proteins.

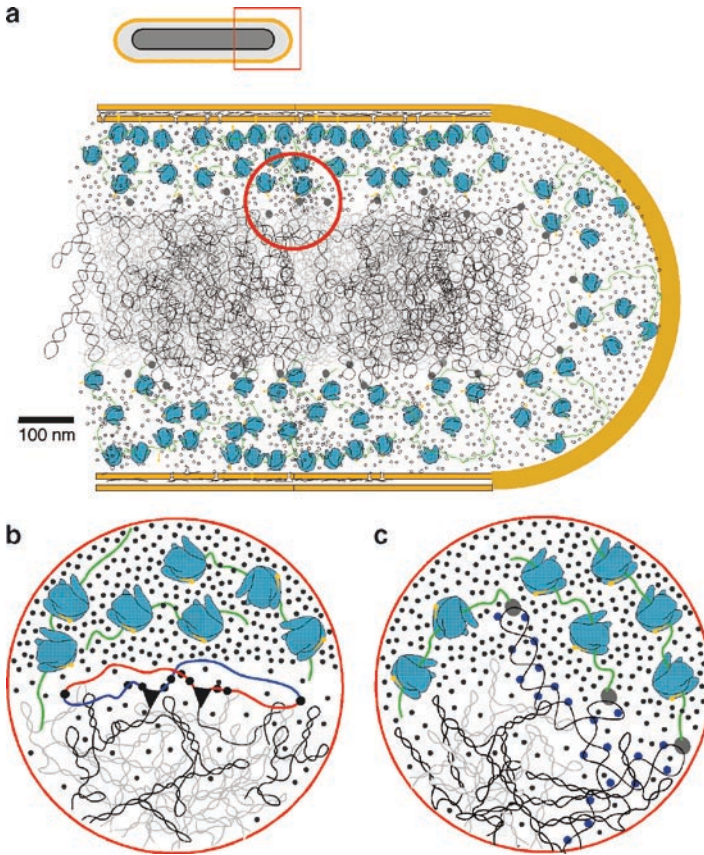


Fig. 5.3 Schematic and hypothetical representations of cytoplasm and nucleoid in a slow-growing *E. coli* cell (cf. Fig. 5.2c). (a) Phase separation between the cytoplasm with numerous soluble proteins and the nucleoid consisting of a congested network of branched supercoils and relatively few proteins. (b) The process of replication is assumed to occur in the border region of the nucleoid. Replisomes are indicated as triangles synthesizing black and grey replichores (cf. Fig. 5.6). (c) Transcription is also assumed to occur in the border region of the nucleoid. DNA-binding proteins (*small circles*) open-up the unconstrained supercoils compacted in the nucleoid center allowing them to become transcribed by RNA polymerases (*large circles*). (Adapted from Fig. 5.1 in Woldringh and Nanninga 2006; see also Fig. 5.4.1 in Goodsell 1993)

The resulting changes in the internal dynamics of the bound segments may bring them to the outer nucleoid border allowing their replication or transcription (as suggested in Fig. 5.3b and c).

Already from early studies onwards the structuring role of DNA binding proteins (and possibly of RNA) has generated the so-called looped-domain model for the bacterial chromosome favoured by many authors (Worcel and Burgi 1972; Trun and Marko 1998; Higgins 1999; Cook 2002; Stavans and Oppenheim 2006; Noom et al. 2007). The model explains the experimental findings of topologically isolated

chromosome domains (see below), but does not help in understanding the degree of nucleoid compaction necessary to pack the nucleoid in the cell. In an extensive study Zimmerman (2006a) determined the role of DNA-associated proteins on nucleoid compaction. He observed that nucleoids isolated from *E. coli* cells by a detergent-polylysine-spermidine procedure (Murphy and Zimmerman 2002) were only partially expanded upon treatment with urea or trypsin. While RNase caused decompaction of the structures, the release of DNA-associated proteins had no effect. This suggests that the importance of DNA binding proteins lies more in their regulatory functions (silencing or modulating gene expression) than in a structural role as a “histone-like” compaction agent or crowding factor; their function thus seems analogous to chromatin remodelling in eukaryotes (Dame 2005).

5.2.4 Isolated Nucleoids

To obtain information about the *in vivo* packing of DNA, many different attempts have been made to open the cells and to characterize the escaped nucleoids. In early studies nucleoids were isolated from *E. coli* cells treated with lysozyme and lysed using detergents in the presence of counterions as NaCl (Worcel and Burgi 1972) or spermidine (Kornberg et al. 1974) to keep the DNA compact. Analysis of the isolated structures by sucrose gradient centrifugation (e.g. Hecht et al. 1977) established the picture that the nucleoid consists of some 50 supercoiled domains separated by barriers that constrain the negative superhelical tension within each domain (for review Higgins 1999). These hypothetical barriers were thought to consist of RNA because the isolated nucleoids unfolded when treated with RNase (Hecht et al. 1977; Murphy and Zimmerman 2000). It has been suggested, however, that the above lysis conditions caused the polyribosomes (but not the RNA polymerases) to dissociate, allowing the naked mRNA to artificially cross-link the escaping nucleoid (Pettijohn 1996; Cook 2002). Nevertheless, genetic studies using various recombinase systems that rely on the superhelicity of the chromosome suggested that the transcriptional activity of the cell could give rise to barriers and the formation of domains with an average size of 10 kbp (for review Deng et al. 2005).

The role of RNA as a cross-linking agent has continued to be investigated using fluorescence microscopy (Zimmerman 2006a) and atomic force microscopy (AFM; Ohniwa et al. 2007 and references therein). Because of the use of detergents in these studies the results have been diverse and are difficult to evaluate. An alternative method that avoids detergents is the osmotic shock of sucrose-protected spheroplasts. From the description in Section 5.2.2 of excluded volume interactions between DNA and soluble proteins it can be expected that the DNA is compacted in the cell like a spring. Indeed, if an *E. coli* spheroplast is broken by osmotic shock, the nucleoid “explodes” out of the spherical ghost and expands into a cloud-like structure (Fig. 5.4; Cunha et al. 2001a, 2005). Although the volume of such a liberated nucleoid (30–100 μm^3) is much larger than its volume in the cell ($\sim 0.1 \mu\text{m}^3$), it is smaller than theoretically expected (166 μm^3 ; Cunha et al. 2001b).

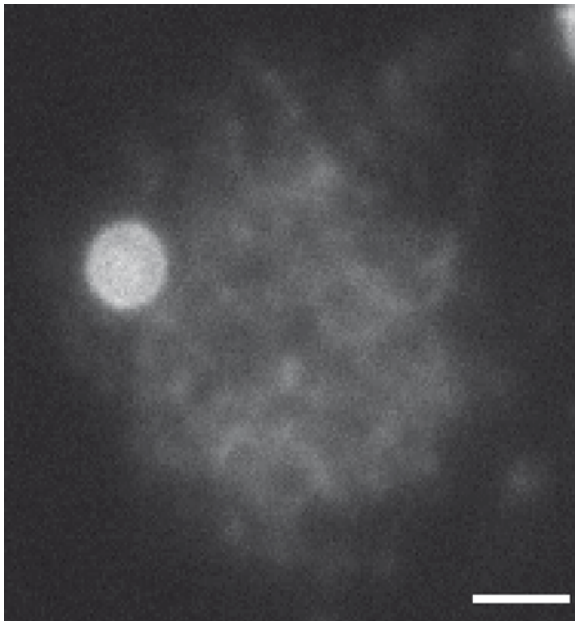


Fig. 5.4 Osmotic shock of an *E. coli* spheroplast stained with DAPI. The nucleoid has “exploded” out of the spheroplast ghost and appears as a cloudlike structure with granules or microdomains which, in the microscope, can be seen to be in Brownian motion. Bar 2 μm

Physical entanglements and cross-links made by DNA binding proteins and possibly by RNA could explain the discrepancy. However, these nucleoids, isolated in the absence of detergents, were not affected by RNase treatment (Cunha et al. 2001a). In contrast, they expanded and dispersed during treatment with proteases (Wegner and Woldringh, unpublished results).

In principle, the measurement of DNA fluctuations by fluorescence correlation spectroscopy (FCS; Romantsov et al. 2007) and the determination of osmotic compressibility as performed by Cunha et al. (2001b) using polyethylene glycol (PEG) as crowding agent are both based on the internal dynamics of the DNA in the nucleoid and should give the same structural information. However they do not. Nucleoids isolated by osmotic shock according to the protocol of Cunha et al. (2001a) consist of irregular clouds (Fig. 5.4) with granular substructures that exhibit Brownian movement. In contrast, Romantsov et al. (2007) obtained rounded, delineated structures, with a volume of about $18 \mu\text{m}^3$ after osmotic shock in the presence of formaldehyde (0.025%) and staining with TOTO-1. From measurements of fluorescence intensity at increasing TOTO-1 concentrations in the same nucleoid, they extracted a value of 50 kbp for the size of a “structural unit”. They estimated its diffusion coefficient to be $6.5 \mu\text{m}^2/\text{s}$, which is much higher than the $0.12 \mu\text{m}^2/\text{s}$ obtained from single-particle tracking of DNA tagged with the Lac-repressor protein, but isolated by another protocol (Cunha et al. 2005). The different results could be ascribed to

variations in growth conditions (see discussion in Romantsov et al. 2007) or to the different method of preparation of nucleoids. Although the discrepancy is presently unresolved, the method of FCS is promising as it can give information on the internal structure of the nucleoid, possibly even in intact cells.

5.2.5 Nucleoid Shape and Transcription

While the role of RNA in determining the domain structures of nucleoids as described in the previous section remains uncertain, its structural role in living cells is more clear. For instance, inhibition of protein synthesis by chloramphenicol has a strong effect on stable RNA synthesis (Shen and Bremer 1977) and dramatically changes the shape of the nucleoid of *E. coli* cells into a spherical body. This was already described in early electron microscope studies (Robinow and Kellenberger 1994) and more recently by phase contrast and fluorescence light microscopy (Zimmerman 2002, 2006a). Inhibition of protein synthesis with chloramphenicol was found to cause a smoothing of the border between nucleoid and cytoplasm in live cells (Fig. 5.5) and, in cells fixed with osmium tetroxide, even a coalescence of segregated nucleoids (van Helvoort et al. 1998). The change in nucleoid shape was observed within a few minutes after chloramphenicol exposure (Zimmerman 2006b), as demonstrated in the timelapse images of live *E. coli* cells with nucleoids tagged with the fluorescent DNA binding protein HupA-RFP (Fig. 5.5). In addition, both the typical dumbbell and W-shapes in fast growing cells (Mason and Powelson 1956; Woldringh et al. 1994; Zimmerman 2002) and the peripheral nucleoids seen in spherical *E. coli* cells (Zaritsky et al. 1999) are abolished upon protein synthesis inhibition (see Fig. 5.4 in Woldringh et al. 1994). These observations suggest that transcriptional activity at the border of the nucleoid (Jin and Cabrera 2006) including transcripts involved in co-transcriptional translation (transertion; Norris and Madsen 1995; Woldringh 2002), can exert forces that shape and position the nucleoid. They formed the basis for the transertion-mediated segregation model (Woldringh 2002).

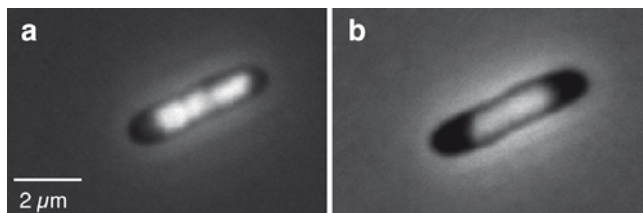


Fig. 5.5 The effect of chloramphenicol on nucleoid structure. **(a)** Live cell from an *E. coli* MC4100/pSACT11 culture rapidly growing in broth medium. Combined phase contrast and fluorescence images. The nucleoid is stained with the fusion protein HupA-RFP expressed from a plasmid. The cell is growing on an agar slab close to a channel cut in the agar. **(b)** Same cell photographed within 1 min after addition of 10 μ M chloramphenicol to the channel

If inhibition of transcription and translation changes the border, shape and position of the nucleoid, does transcription also influence the internal organization of the DNA within the nucleoid? Recent studies that describe short- and long-range correlations in co-expression of genes in *E. coli* seem relevant to this. As the transcription of a certain gene has an effect on the superhelicity of the surrounding DNA (Higgins 1999) it may induce the transcription of its adjacent genes leading to a short range (up to 16 kb) spatial correlation of expressed genes. Surprisingly, however, measurement of mRNA-abundance genes occurring in micro-array studies also indicated long-range correlations corresponding to distances of ~100 and ~600 kb (Jeong et al. 2004; Carpentier et al. 2005). Such correlations between co-expressed genes that have a mere spatial but no physiological relationship, led to the suggestion that the nucleoid is organized in the form of two types of spirals or loops consisting of coiled, unexpressed and uncoiled, expressed DNA (cf. Fig. 5.6 in Carpentier et al. 2005; Riva et al. 2008). Future experiments with cells in which mRNA can be detected at multiple chromosomal reporter sites using fluorescent RNA-binding proteins (e.g. Golding and Cox 2004) could shed light on a possible long-range transcriptional organization of the nucleoid.

5.3 Chromosome Segregation

5.3.1 Dynamics of DNA Spots During Segregation

It has often been reported (e.g. Ghosh et al. 2006) that DNA regions move abruptly and faster than cell elongation and that such observations are indicative of the existence of a directed driving force and a dedicated segregation mechanism similar to that seen in eukaryotic mitosis (cf. Fig. 5.1d). This view is generally based on reports of DNA dynamics that show variable results; they are often based on few observations and are not always sufficiently documented with respect to the method of measuring or to the cell's physiology.

Table 5.1 gives an overview of published data on the dynamics of DNA regions either expressed as diffusion coefficient or as spot displacement in time. In the experiments of Elmore et al. (2005) on the movement of *oriC*-GFP spots in *E. coli* cells growing in rich medium (multifork replication; Fig. 5.1b), it was found that after correction for cell elongation the spots only showed erratic, Brownian trajectories. Their movement was characterized by a Gaussian distribution of step sizes and a very low effective diffusion coefficient. Correction for cell elongation (0.07 $\mu\text{m}/\text{min}$) was performed by assuming that the cell can be represented by a homogeneous elastic body that is stretched by affine deformation (i.e. by maintaining the relative distance between all structures in the body). The conclusion was that the drift caused by such an affine deformation is sufficient for spot segregation along the length axis of the cell. In other words, no additional drift, for instance caused by an active mechanism, seemed to be necessary for segregation. The values for the diffusion coefficient in the long and the short cell axis ($D = 4.3 \times 10^{-5}$ and $2.9 \times 10^{-5} \mu\text{m}^2/\text{s}$, respectively).

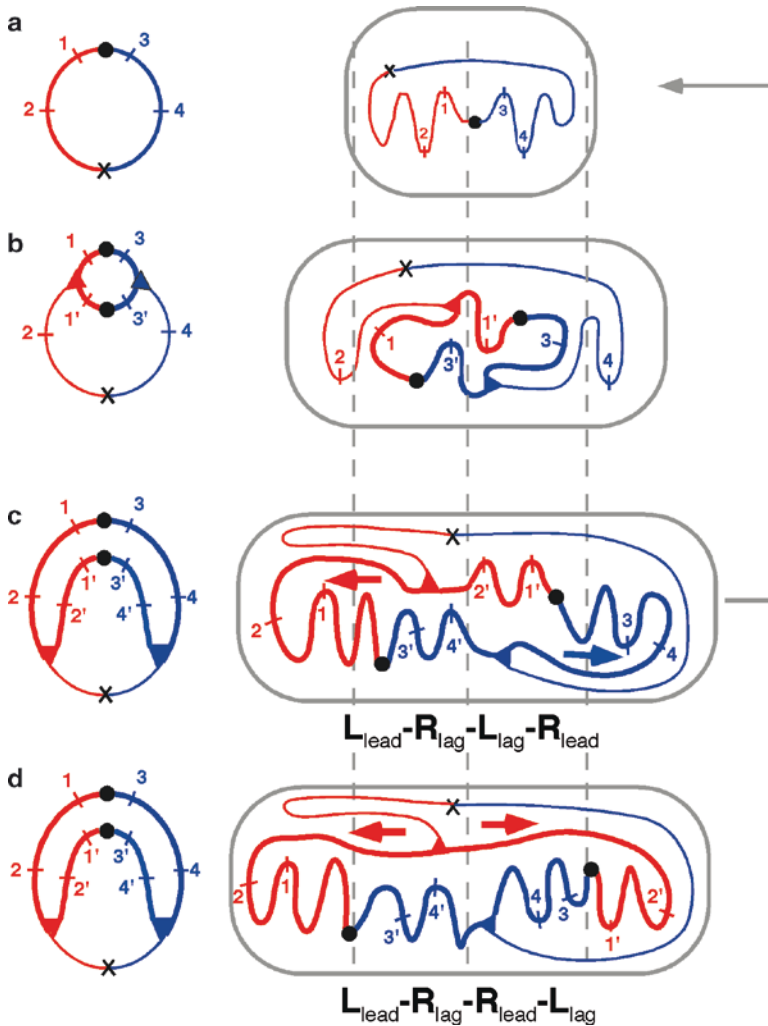


Fig. 5.6 Ordering of the two chromosome arms (black and grey) in *E. coli*. Compare with Fig. 5.6 in Nielsen et al. (2006b) and Fig. 5.2 in Reyes-Lamothe et al. (2008a). *OriC* is drawn as a black circle and *terC* as a X. The loci 1, 2 and 3, 4 are on the leading strands, while 1', 2', 3' and 4' are on the lagging strands. The two replisomes are represented by black and grey triangles. **(a)** Non-replicating chromosome. The two replichores are lying in two halves of the nucleoid with *oriC* in the middle and *terC* at the side of the new cell pole. **(b)** Upon initiation, unreplicated DNA from either chromosome arm is reeled in while the DNA of the daughter strands, each consisting of a black and a grey arm, move away from each other. **(c)** An asymmetric ordering (L-R-L-R) is obtained when each daughter strand has either a fast-moving (markers 1, 2 and 3, 4) or a slow moving arm (markers 1', 2' and 3', 4'). The fast moving arms indicated by arrows, move past the origins. **(d)** Symmetric ordering of chromosome arms (LRLR or RLLR) is obtained when in each daughter strand the same chromosome arm is either fast-moving (markers 1, 2, and 1', 2') or slow-moving (markers 3', 4' and 3, 4)

Table 5.1 Comparison of diffusion coefficients and spot displacements of DNA loci and the rate of cell elongation in different bacteria

DNA or bacterial strain	Diffusion coefficient (μm ² /s)	Δ Spot position/ time ^a (μm/min)	Rate of cell elongation (μm/min)	Reference
<i>E. coli</i>		~10× Elongation rate		Gordon and Wright 2000
<i>B. subtilis</i>		0.17–0.27	0.02	Webb et al. 1998
<i>E. coli</i>	4.3 × 10 ⁻⁵ (stepsize 110 nm)		0.07	Elmore et al. 2005
Long axis				
Short axis	3 × 10 ⁻⁵			
<i>C. crescentus</i>		0.1–0.4	0.006	Viollier et al. 2004
		0.3		Toro et al. 2008
<i>V. cholerae</i>	2–4 × 10 ⁻⁴	0.06	0.02	Fiebig et al. 2006
Long axis	(stepsize 250 nm)			
Short axis	1–3 × 10 ⁻⁴			
<i>E. coli</i>	5 × 10 ⁻⁴	0.4		Reyes-Lamothe et al. 2008a
Replisome	(stepsize 100 nm)			
DNA spot L3	5 × 10 ⁻⁵			
<i>E. coli</i>		0.2	0.015	Wang et al. 2005
<i>ter</i> -marker				

^aOften expressed as rate or speed of movement. However, speed in μm/min is not well defined in the case of random, Brownian diffusion (see Berg 1993).

Elmore et al. (2005) suggest that the *oriC*-GFP spots are virtually immobile and exhibit subdiffusive behavior within a confined region (i.e. less motion than expected from free diffusion). The size of this region or “cage” can be estimated by applying the formula for the mean square displacement (*MSD*) of a particle diffusing in two dimensions, *MSD* = 2*tD* (Berg 1993). With a diffusion coefficient of *D* = 3 × 10⁻⁵ μm²/s, a DNA spot could “explore” within a time interval of *t* = 60 min a region with a diameter of about 0.5 μm, roughly corresponding to the width of the nucleoid (Elmore et al. 2005). A similar size for the caged domain of a DNA locus was estimated by Fiebig et al. (2006) for the smaller *V. cholerae* cell.

Low diffusion coefficients for DNA segments have now been reported by several groups (e.g. in the order of 10⁻⁴ μm²/s; see references in Table 5.1). This restricted movement or virtual immobility of the DNA can be understood when considering the nucleoid (see Fig. 5.3) as a highly congested, visco-elastic network (Woldringh and Odijk 1999) and could explain why during replication loci become and remain positioned in the cell in the same order as in which they occur on the genetic map (Viollier et al. 2004; Nielsen et al. 2006a), without having to invoke putative capture sites and anchoring proteins. Such proteins have nevertheless been proposed to exist in *Bacillus subtilis* cells (Berkmen and Grossman 2007) because positioning of the origin at the cell quarters was found to be independent of where replication started on the chromosome (at 0° or 257°). Moreover, in sporulating *B. subtilis* cells, origins are found at opposite ends of the nucleoid probably captured and tethered near the cell poles with the help of anchoring proteins (Errington et al. 2005).

The drift caused by affine deformation could be the result of two biological processes: (i) protein and cell envelope synthesis (transcription, translation and translocation) that cause elongation of the cell body, and (ii) DNA synthesis (replication, segregation) that causes extension of the visco-elastic nucleoid. While such a drift seems to be adequate for segregation over the relatively short distances in the nucleoids of *E. coli* cells growing in rich medium (Elmore et al. 2005), it seems to be insufficient in *C. crescentus* (Viollier et al. 2004; see Fig. 5.3 in Woldringh and Nanninga 2006) and in *V. cholerae* (Fiebig et al. 2006). In these cells the origin is positioned at the cell pole and requires movement over a larger distance. Thus, a special mechanism seems to be necessary for the initial separation of the spots and possibly for its further displacement in *Vibrio* and *Caulobacter*. However, it will be argued below that such a mechanism could be furnished by volume exclusion effects and thus by entropic, repulsive forces between newly replicated DNA strands as proposed by Jun and Mulder (2006).

5.3.2 Ordering of Chromosome Arms (Replichores)

Two methods have been applied to visualize genetic loci in bacteria: fluorescence in situ hybridization (FISH) and the fluorescent repressor protein/operator system technique (FROS). With the first technique it was shown that bacterial chromosomes are organized in the fixed cells in an order that reflects their position on the genetic map with the origin and terminus occupying polar positions (Niki et al. 2000; Roos et al. 2001; Bates and Kleckner 2005). This was most clearly demonstrated in *C. crescentus* by Viollier et al. (2004). It was therefore a surprise that visualization of loci with FROS in living cells revealed a different organization: in *E. coli* the chromosome arms were found to occur in the two halves of the nucleoid with the origin in the center (Fig. 5.6a). This was first discovered in the laboratory of D. Sherratt (Wang et al. 2005, 2006) and confirmed in the laboratories of S. Austin and F. Hansen (Nielsen et al. 2006a,b). Apparently, in *C. crescentus*, newly replicated loci on both chromosome arms move at the same velocity and become sequentially placed at the inside of already replicated loci. While one origin remains at the stalk cell pole the other origin is pushed to and kept at the polar tip of the newly synthesized nucleoid with both chromosome arms moving up side by side (see Fig. 5.1 in Wang et al. 2006).

On the other hand, in *E. coli*, newly replicated loci on different chromosome arms seem to move in opposite directions. Moreover, for the origins to end up in the middle of the newly synthesized nucleoids, one chromosome arm of each daughter strand has to move faster than the origin. This is illustrated in Fig. 5.6, where it is assumed that the replisomes are independently moving structures that reel in and thereby move to that half of the nucleoid of which they replicate the DNA (Fig. 5.6b). Newly replicated DNA is assumed to become sequentially layered on either side of the origins (cf. Fig. 5.2 in Reyes-Lamothe et al. 2008b). Progressive replication of the chromosome can lead to a different ordering pattern

of the genetic loci (indicated by the numbers in Fig. 5.6), depending on which chromosome arm moves faster. For instance, an asymmetric ordering is obtained if the leading strand moves faster than the lagging strand ($L_{lead}-R_{lag}-L_{lag}-R_{lead}$; Fig. 5.6c). A symmetric ordering (either $L_{lead}-R_{lag}-R_{lead}-L_{lag}$ or $R_{lag}-L_{lead}-L_{lag}-R_{lead}$) is obtained if both the leading and lagging strand can move faster than those of the other chromosome arm (Fig. 5.6d). If there were no preference and one strand overtakes the origin by chance, we would expect to see the L-R-L-R pattern in 50% of the cells and each of the two symmetric patterns in 25% of the cells.

Differences in the velocity of strand movement have been ascribed to differential gene expression on the leading and lagging strand (Woldringh and Nanninga 2006). Because immediately after replication the lagging strand is involved in the ligation of Okazaki fragments, an imbalance may develop in the availability of genes for transcription. This may induce a differential gene expression between the two genetically identical daughter strands resulting in different mobilities of the leading and lagging strands of the two chromosome arms (Wang et al. 2005). Another possibility, suggested by White et al. (2008) is that the looped configuration of the lagging strand in the replisome would slow down its motion.

In *E. coli*, however, the ordering appears to show quite some variation, probably related to the growth rate of the cell or to specific proteins like MukBEF (Danilova et al. 2007). Nielsen et al. (2006b) estimated that in 20–25% of the cells the replisomes showed a symmetric ordering of LRRL or RLLR (Fig. 5.6d), suggesting that for either daughter both the leading and the lagging strand can move faster. In this case transcription does not play a role in distinguishing the strands (see also Wang et al. 2006). Below it will be discussed further what role transcription plays in DNA movement and replisome ordering.

5.3.3 Replisome Movement and Cohesion

The positions of replisomes as drawn in Fig. 5.6 are based on the timelapse studies of A. Wright (Fig. 5.7) and Reyes-Lamothe et al. (2008b) that indicate that replisomes tagged with Ssb-CFP can assemble at the origin irrespective of its location in the

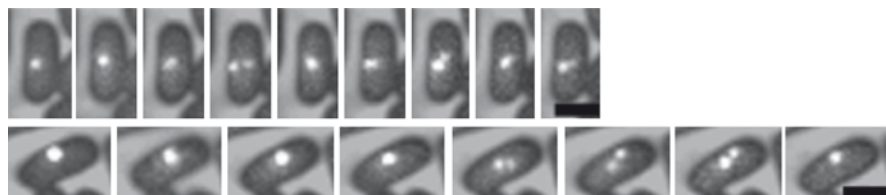


Fig. 5.7 Two timelapse sequences of replisomes on the chromosome of slow growing *E. coli* cells (strain MG1655) visualized by fluorescent single-stranded binding protein (SSB-GFP). The spots first move independently of each other, then converge at the end of the sequences during termination of replication. Photographs by Andrew Wright. Time interval 2 min. Magnification bar 1 μ m

nucleoid and move from there independently from each other. They can come together or drift apart, but return to midcell at termination where they will disappear.

According to the analysis of the *MSD*-plots against time (Reyes-Lamothe et al. 2008b), the dynamics of replisomes in the long axis of the cell is about 10x faster than that of genetic loci (Table 5.1). Is this because the replisomes are always located in the border region of the nucleoid (Fig. 5.3b where they are being moved because of the separation of the newly replicated daughter strands? As noted by the authors more accurate single particle tracking measurements are necessary to distinguish the different dynamics of loci in the process of replication (Reyes-Lamothe et al. 2008b; see also discussion on the accuracy of these analyses in Elmore et al. 2005). The timelapse observations of live cells seem to contradict previous studies on fixed cells (Bates and Kleckner 2005; den Blaauwen et al. 2006) that suggest the existence of replication factories. The observation of replisomes staying together could result from the mere fact that they are connected to the bulk of unreplicated DNA in the nucleoid (cf. Fig. 5.3b).

Another controversial aspect of bacterial DNA segregation has been the cohesion of newly replicated daughter strands, that would resemble the pairing of sister chromatids by the eukaryotic cohesion complex (Nasmyth 2002; cf. Fig. 5.1c). Although a function for bacterial cohesion has not been proposed, a large delay in the separation of fluorescent origin spots has been determined in many studies (Bates and Kleckner 2005; Nielsen et al. 2006a; Adachi et al. 2008), while in other studies only a small or no delay was found (Roos et al. 2001; Li et al. 2002; Elmore et al. 2005; Jensen 2006; Nielsen et al. 2007).

When comparing these different results it should be realized that the estimated delay is the outcome of two different techniques that both have their limitations: (i) Flow cytometry of a culture treated with rifampicin and cephalixin (for ~4 h) to allow run-off DNA synthesis while inhibiting new initiations and cell divisions, respectively. For a slow-growing population ($T_d > 60$ min) one sees a bimodal distribution of the amount of DNA per cell. The two peaks represent cells with 1 and 2 chromosome equivalents (or origins). From the proportion of cells with 1 origin the age at initiation of replication is calculated (a_i ; see formula below). The smaller this proportion is, the earlier initiation occurs in the cell cycle. (ii) Image cytometry of a usually fixed population of cells prepared for fluorescence microscopy. Scoring the fluorescent *oriC*-spots in the cells (visualized by either FISH or FROS) gives the proportion of cells with 1 spot, from which the age of origin duplication or segregation is determined. Also, the smaller this proportion is, the earlier segregation occurs.

The age at initiation or at segregation (a_i) is calculated by making use of the age distribution function ($n(a) = n_0 \cdot 2^{-a/t}$, in which a is age and t the generation time). After normalization and integration of $n(a)$ from age = 0 to age a_i at initiation or segregation, we obtain the formula $a_i = -(t/\ln 2) \cdot \ln(1 - 0.5P)$, in which P is the proportion of cells with one *oriC* copy or one fluorescent spot, from which we can calculate a_i . A prerequisite for the application of the age distribution function is that the culture should be in a steady state of growth. This, however, is usually not documented. In many studies the proportion of cells with one *oriC* copy as determined with the flow cytometer is smaller than that of microscopically scored cells with one *oriC* spot, suggesting early

initiation and late segregation of spots in the cell cycle and thus a delay in segregation; this has often been interpreted to reflect cohesion.

As discussed by Reyes-Lamothe et al. (2008a) (see also discussions in Jensen 2006; Woldringh and Nanninga 2006) the high proportion of cells with only one *oriC* spot obtained with microscopic image cytometry (suggesting late segregation) could be caused by low labeling efficiency in FISH experiments (Nanninga et al. 2002) or by sticking of spots due to overproduction of the fusion proteins in FROS-experiments (see also Nielsen et al. 2006a). A higher proportion of one *oriC* spots is also expected because duplicated spots are not immediately resolved optically because of the limited resolution of the light microscope (see Fig. 5.2 in Woldringh and Nanninga 2006) and because of the replication-induced formation of precatenanes that have to be resolved by topoisomerase IV (Wang et al. 2008). Timelapse experiments of slow-growing cells from which the cell cycle parameters (C and D periods) are well known and in which overexpression of fluorescent repressor is prevented seem to have ended the existence of prokaryotic cohesion as a dedicated mechanism (Valens et al. 2004; Elmore et al. 2005; Nielsen et al. 2007; Reyes-Lamothe et al. 2008b). As discussed above (Fig. 5.1c) cohesion is functional in eukaryotes where it is a prerequisite for mitotic segregation (Fig. 5.1d).

5.3.4 Active Mechanisms for Chromosome Segregation

In the literature there have been suggested two mechanisms that could play an active role in segregation: assembly of (cytoskeletal) protein (Gitai et al. 2005; Toro et al. 2008) and the process of transection (Norris and Madsen 1995; Woldringh 2002). For *B. subtilis*, *C. crescentus* and *E. coli* the shape-determining, actin-like protein MreB has been implied to be a key protein for segregation. However, this is no longer plausible as Karczmarek et al. (2007) found no segregation defect in an *E. coli mreBCD* deletion strain. Their careful comparison of the rod-to-sphere transition upon inhibition of MreB with A22 and of penicillin-binding protein 2 with mecillinam, showed how DNA segregation can proceed in both types of spherical cells in the absence of functional MreB proteins.

Another proposed mechanism is transection-mediated segregation (Norris and Madsen 1995; Woldringh 2002). It was discussed above that transcription plays a role in shaping the nucleoid (Fig. 5.5), but does it also play a role in the separation of newly replicated DNA strands as previously suggested? Observations by Nielsen et al. (2006a), indicated that when transcription is inhibited with rifampicin allowing run-off DNA synthesis, replicated DNA loci do not stay together but become distributed to the two cell halves (see Fig. 5.7 in Nielsen et al. 2006a). This experiment has been confirmed with an *E. coli* strain containing a differently coloured fluorescent tag on each of the two chromosome arms (Hansen and Woldringh, unpublished results). Upon treatment of exponentially growing cells with rifampicin and cephalixin for 3 h the average number of spots per cell increased from 1.2 to 1.6 (number of cells examined > 400), indicative of run-off DNA synthesis and

initial spot separation. If transcription were the driving force for segregation and replicore ordering (see Fig. 5.6), the rifampicin-treated population would be expected to show many spots in pairs and thus an ordering like LLRR. However, this was not observed. It should be noted that the percentage of cells showing 4 fully separated spots was only 22% in the control and 43% in the rifampicin-treated population. In both cases about half of the cells showed an asymmetric ordering of LRLR or RLRL, one quarter the LRRL and another quarter the RLLR-ordering.

These preliminary experiments indicate that spots replicated in the absence of transcription, can become separated over relatively large distances and become positioned into both cell halves. The ordering suggests a random choice of either the leading or the lagging strand moving faster, contrary to the observations of White et al. (2008) that it is the leading strand that moves faster. Future (timelapse) experiments with two fluorescent tags in cells grown under steady state conditions, will have to establish whether asymmetric replicore ordering can be more pronounced and if so, what the underlying mechanism is that distinguishes the strands. In the next section I will consider whether segregation could take place without an active mechanism.

5.3.5 Segregation by Entropic Forces

In Fig. 5.3 a rather chaotic picture has been drawn of the huge, branched supercoil in the *E. coli* nucleoid. The drawing is based on experimental and theoretical considerations of Odijk (Odijk 1998; Cunha et al. 2001b) in which the superhelical chromosome with a supercoil contour length of 620 μm can be represented by a chain of 4,000 Kuhn segments that have a length of 158 nm and form trifunctional branches with a density of about 0.6 branch/kbp. Figure 5.3a shows the packing of these segments within the nucleoid without any indication of substructures, cross-links or domains. On the other hand, as discussed above, several observations suggest that the branched supercoil is subdivided. For instance, the 10 kbp domains described by Deng et al. (2005), the presence of some 360 crosslinks deduced from the free energy estimate of isolated and compressed nucleoids (Cunha et al. 2001b), the structural units of 50 kbp measured by Romantsov et al. (2007) and the granularity exhibited by isolated nucleoids (cf. Fig. 5.4) all seem to indicate the presence of nucleoid substructures.

de Gennes (1979) described that if a linear polymer were squeezed in a narrow tube the chain would behave like a sequence of non-overlapping “blobs” (see also Reisner et al. 2005). It is therefore tempting to assume that the branched supercoil, highly confined within the nucleoid, will likewise form blobs. However, the fact that we are dealing with a branched supercoiled structure instead of linear DNA makes it much more difficult to make a physical prediction of its confined configuration. Nevertheless it has been assumed by several authors that the supercoiled segments will become compacted into a string of “blobs” (Jun and Mulder 2006; Fan et al. 2007).

In Fig. 5.8a this possibility has been depicted for an arbitrary stretch of ten adjacent, branched Kuhn segments with a supercoil length of $1.6\ \mu\text{m}$ that are confined in a “blob” with a diameter of $80\ \text{nm}$ (volume $\sim 0.0003\ \mu\text{m}^3$; size $\sim 12\ \text{kbp}$). The resulting degree of compaction is the same as that for the whole chromosome (supercoil length of $620\ \mu\text{m}$) in the nucleoid volume ($\sim 0.1\ \mu\text{m}^3$). In Fig. 5.8b a sequence of several blobs is indicated as they might occur within the nucleoid. It should be emphasized that these “blobs” do not represent permanent, physical structures, but merely reflect regions within the network in which fluctuations of

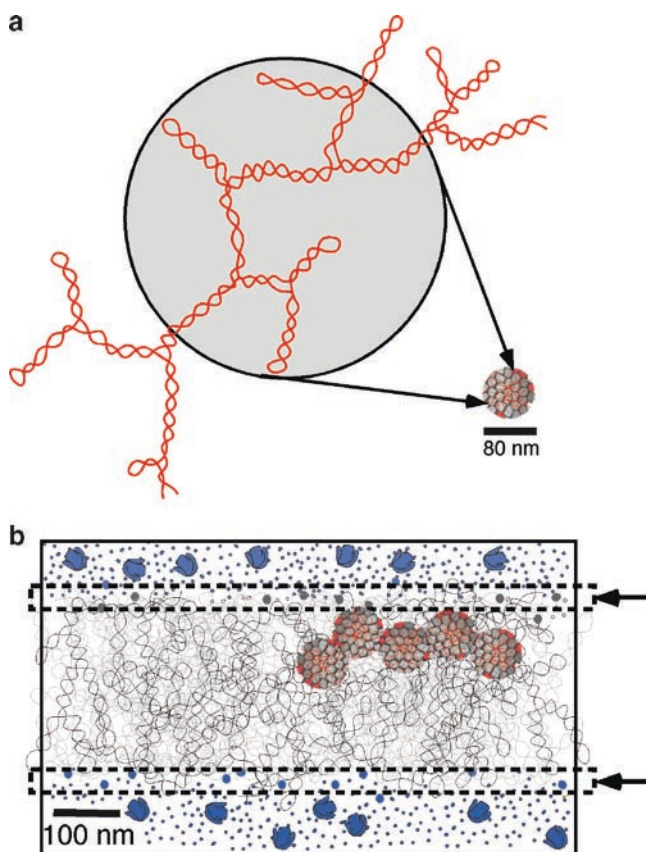


Fig. 5.8 DNA compaction in the nucleoid. **(a)** An extended stretch of the branched supercoiled chromosome consisting of Kuhn segments with a persistence length of $158\ \text{nm}$. Within the large circle 10 Kuhn segments are drawn in 2-D; they represent a supercoil length of $1.6\ \mu\text{m}$ and $\sim 12\ \text{kbp}$ of DNA. When assuming the same degree of compaction as for the whole chromosome, the 10 Kuhn segments have to be confined within a 3-D globule (“blob”) with a diameter of $80\ \text{nm}$ (volume $0.0003\ \mu\text{m}^3$). Arrows indicate the necessary compaction. **(b)** Within the chaotic representation of the nucleoid (cf. Fig. 5.3) a string of “blobs” as defined in **(a)** has been indicated. Arrows mark the border region or concentric shell of the nucleoid

the polymer segments are correlated. The regions, transiently determined by crosslinks, may vary in size and are defined by the average correlation length. (The possible role of proteins in determining these regions or crosslinks can be considered as incorporated in the “blob” model.)

Given such a string model of the nucleoid, how can we envisage its replication and segregation? According to Jun and Mulder (2006) the same excluded volume interactions that cause the nucleoid to “explode” out of a spheroplast (Fig. 5.4) also cause daughter strands to resist their overlapping and to segregate spontaneously. They propose that the free energy of the system will decrease when each of the interconnective chains can move into a separate blob rather than intermingle and collide with the other chain. However, to allow this free-energy driven separation of chains to occur within the biological time scale of a bacterial cell cycle they had to make several assumptions. First, a physical mechanism cannot distinguish between daughter strands consisting either of two different replichores with the origin in the middle or of two identical replichores with the replisome in the middle (see Fig. 5.6b). To allow the daughter strands to drift apart rather than the replichores, they had to assume that replisomes stick together for some initial period; this is feasible because of the connection of the replisomes with the bulk of nucleoid DNA from which they reel in the unreplicated strands. Second, they had to assume a much higher diffusion coefficient for segregating DNA than the experimentally obtained values (see Table 5.1). They therefore proposed a so-called “concentric shell” model in which the newly replicated strands occur in the border region between nucleoid and cytoplasm (Fig. 5.3b and c). In this two-dimensional concentric shell (arrows Fig. 5.8b) the entropic, repulsive forces between newly replicated chains would be much stronger causing a higher diffusion coefficient and assuring their separation.

An interesting aspect of the chromosome model of Jun and Mulder (2006) is that plasmids are predicted not to become segregated by entropy because of their small size and insufficient concentration. They will therefore need a more dedicated mechanism like the Par-system described by Garner et al. (2004). A similar system (*parABS*) has been proposed to be required for initial separation of replicated DNA in *Caulobacter* (Toro et al. 2008).

With their assumptions, Jun and Mulder (2006) could simulate the segregation process in *E. coli* as described by Bates and Kleckner (2005) and in *C. crescentus* as found by Viollier et al. (2004). Also based on excluded-volume interactions and a series of assumptions concerning the properties of interlinked chromosomal domains, Fan et al. (2007) simulated the segregation and even the replicore ordering in the *E. coli* nucleoid. What these simulation studies based on physical mechanisms seem to indicate is that with adequate assumptions bacterial DNA segregation may not need an active, dedicated mechanism. However, hidden behind the assumptions may lie biological mechanisms. For instance, underlying the concentric shell model (Jun and Mulder 2006) with increased dynamics for DNA strands, could be a changed conformation of newly synthesized DNA as suggested in Fig. 5.3b.

In a recent fluorescence microscopic study of living *B. subtilis* cells, early and late replicated DNA could be distinguished by the incorporation of fluorescent nucleotides with two colors (Berlitzky et al. 2008). The deconvolved images

suggest that the newly synthesized DNA moves over the surface of the unreplicated nucleoid toward the poles where it accumulates. Further studies have to establish whether these observations support the idea of a concentric shell.

5.4 Conclusions and Questions

During the bacterial cell cycle DNA strands become segregated soon after their replication. Such a mechanism ensures that re-initiations can take place during multifork replication and rapid growth. This phenomenon of overlapping replication cycles makes the segregation mechanism fundamentally different from that in eukaryotes where replicated strands remain together by cohesion until mitosis (Fig. 5.1).

Non-specific, excluded-volume interactions between DNA and soluble proteins are the main cause of a phase separation (Figs. 5.2 and 5.3) between nucleoid and cytoplasm (Odijk 1998; Woldringh and Odijk 1999). Nevertheless, the shape of the nucleoid can be altered by the process of transcription as indicated by the changes observed upon inhibition of RNA or protein synthesis with either rifampicin or chloramphenicol (Fig. 5.5). Although in the nucleoid the DNA is present as a dense, visco-elastic network of branched supercoils, it seems possible that the nucleoid contains substructures in the form of a string of globular “blobs” (Fig. 5.8).

In the simultaneous replication/segregation process of different bacteria the movement of chromosome arms seems to differ. While in *Caulobacter* and *Vibrio* the two replichores appear to move together, they segregate at different velocities in *E. coli* giving rise to a positioning of each replichore in either half of the nucleoid with the origin in between (Fig. 5.6). To explain this ordering of replichores Wang et al. (2006) proposed that the replichores are directed by the independent movement of the replisomes into opposite cell halves, where they layer the DNA sequentially on both sides of the origin. Likewise, Nielsen et al. (2006b) proposed a “rearrangement mechanism” which moves one replichore past and on the outside of the origin (Fig. 5.6). White et al. (2008) found indications that it is the replichore of the leading strand that moves faster. Whatever the mechanism ensuring this differential movement and ordering of the replichores, it is remarkable that it also occurs during run-off DNA synthesis in the absence of growth (Nielsen et al. 2006a; Reyes-Lamothe et al. 2008a; Hansen and Woldringh, unpublished). Whereas these observations falsify the transertion-mediated segregation model (Woldringh 2002), they strengthen the possibility that entropy is the main driving force in segregation.

However, the diffusion coefficient of DNA loci is very low (Table 5.1) and the dynamics of duplicated origins in fast growing cells suggest that after correction for cell elongation the spots are virtually immobile (Elmore et al. 2005). This prompted Jun and Mulder (2006) to make the assumption of a nucleoid border region or “concentric shell” in which the dynamics of DNA strands is (transiently) higher than within the congested nucleoid. A fast separation of newly replicated DNA in the

outer nucleoid border could initially move the strands far apart by entropic forces before they become again embedded inside the nucleoid. This model allowed simulation of the segregation patterns as observed in *E. coli* and *C. crescentus* (Jun and Mulder 2006). The low diffusion coefficient of the bulk of the DNA could explain why replicated loci keep their position within the nucleoid in the order in which they have been replicated (Viollier et al. 2004; Wang et al. 2005; Nielsen et al. 2006a).

Although the above description gives a quite coherent picture of nucleoid organization many questions remain: (i) Does the nucleoid contain substructures (domains or “blobs”) and if so, what is their size and stability and are proteins involved in their behavior? (ii) Is there a difference in dynamics and positioning between newly replicated DNA and unreplicated or earlier replicated DNA, also during run-off DNA synthesis? (iii) Does the nucleoid border represent a narrow region (“concentric shell”) where replication, segregation and transcription occur? (iv) What mechanism determines the differential dynamics of leading and lagging strands? (v) Does transcription influence the dynamics of chromosomal loci and does it induce long-range correlations in gene expression? (vi) How are sister nucleoids separated before cell division against the excluded volume interactions that cause the phase separation?

The analysis of timelapse and snapshot images of strains that contain multiple fluorescent tags would greatly improve when cell cultures are grown under defined physiological conditions (i.e. in steady state of growth). With a combination of new labeling techniques (e.g. Berlatzky et al. 2008) and improved fluorescence detection methods (e.g. fluorescence correlation spectroscopy and microscopy) future experiments will give better insight in the dynamics and positioning of DNA loci and replisomes during bacterial segregation. When the new data could also be interpreted and understood in terms of polymer physics, some of the questions above might become answered.

Acknowledgements This work is supported by the Stichting voor Fundamenteel Onderzoek der Materie (FOM) and the Stichting Aard- en Levenswetenschappen (ALW), both subsidiaries of the Netherlands Organization for Scientific research (NWO). I am grateful to Elaine Newman for critical reading of the manuscript and constructive comments and to Suckjoon Jun for helpful discussions. I thank Nanne Nanninga for his support and Theo Odijk for introducing me to polymer physics. I especially thank Norbert Vischer for our long-standing cooperation and for his help with image cytometric analyses (<http://simon.bio.uva.nl/ObjectJ/index.html>). I also thank Shun Adachi for performing the “cohesion experiment” mentioned in this work, Andrew Wright for discussions and for unpublished work (Fig. 5.7), Flemming Hansen for collaboration with the run-off DNA replication experiment, Lucy Shapiro for the *Caulobacter* strain (Fig. 5.2) and Tanneke den Blaauwen and Anna Wegner for discussions and results on isolated nucleoids.

References

- Adachi S, Fukushima T, Hiraga S (2008) Dynamic events of sister chromosomes in the cell cycle of *Escherichia coli*. *Genes Cells* 13:181–197
- Azam TA, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181:6361–6370
- Bates D, Kleckner N (2005) Chromosome and replisome dynamics in *E. coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. *Cell* 121:899–911

- Bayer ME (1991) Zones of membrane adhesion in the cryofixed envelope of *Escherichia coli*. J Struct Biol 107:268–280
- Berg HC (1993) Random walks in biology, 2nd edn. Princeton University Press, Princeton, NJ
- Berkmen MB, Grossman AD (2007) Subcellular positioning of the origin region of the *Bacillus subtilis* chromosome is independent of sequences within *oriC*, the site of replication initiation, and the replication initiator DnaA. Mol Microbiol 63:150–165
- Berlatzky IA, Rouvinski A, Ben-Yehuda S (2008) Spatial organization of a replicating bacterial chromosome. Proc Natl Acad Sci USA 105:14136–14140
- Borgnia MJ, Subramaniam S, Milne JLS (2008) Three-dimensional imaging of the highly bent architecture of *Bdellovibrio bacteriovorus* by using cryo-electron tomography. J Bacteriol 190:2588–2596
- Carpentier AS, Torresani B, Grossmann A, Hénaut A (2005) Decoding the nucleoid organization of *Bacillus subtilis* and *Escherichia coli* through gene expression data. BMC Genom 6:84. doi:10.1186/1471-2164-6-84
- Cook PR (2002) Predicting three-dimensional genome structure from transcriptional activity. Nature Gen 32:347–352
- Cunha S, Odijk T, Sueleymanoglu E, Woldringh CL (2001a) Isolation of the *Escherichia coli* nucleoid. Biochimie 83:149–155
- Cunha S, Woldringh CL, Odijk T (2001b) Polymer-mediated compaction and internal dynamics of isolated *Escherichia coli* nucleoids. J Struct Biol 136:53–66
- Cunha S, Woldringh CL, Odijk T (2005) Restricted diffusion within the isolated *Escherichia coli* nucleoid. J Struct Biol 150:226–232
- Dame RT (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Mol Microbiol 56:858–870
- Danilova O, Reyes-Lamothé R, Pinskaya M, Sherratt D, Possoz Ch (2007) MukB colocalizes with the *oriC* region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. Mol Microbiol 65:1485–1492
- de Gennes PG (1979) Scaling concepts in polymer physics. Cornell University Press, Ithaca, NY
- den Blaauwen T, Aarsman MEG, Wheeler LJ, Nanninga N (2006) Pre-replication assembly of *E. coli* replisome components. Mol Microbiol 62:695–708
- Deng S, Stein RA, Higgins NP (2005) Organization of supercoil domains and their reorganization by transcription. Mol Microbiol 57:1511–1521
- Donachie WD, Addinall S, Begg KJ (1995) Cell shape and chromosome partition in prokaryotes or, why *E. coli* is rod-shaped and haploid. Bioessays 17:569–576
- Elmore S, Müller M, Vischer NOE, Odijk Th, Woldringh CL (2005) Single-particle tracking of *oriC*-GFP fluorescent spots during chromosome segregation in *Escherichia coli*. J Struct Biol 151:275–287
- Eltsov M, Dubochet J (2005) Fine structure of the *Deinococcus radiodurans* nucleoid revealed by cryoelectron microscopy of vitreous sections. J Bacteriol 187:8047–8054
- Errington J, Murray H, Wu LJ (2005) Diversity and redundancy in bacterial chromosome mechanisms. Philos Trans Roy Soc B 360:497–505
- Fan J, Tuncay K, Ortoleva PJ (2007) Chromosome segregation in *Escherichia coli* division: a free energy-driven string model. Comput Biol Chem 31:257–264
- Fiebig A, Keren K, Theriot JA (2006) Fine-scale time-lapse analysis of the biphasic, dynamic behaviour of the two *Vibrio cholerae* chromosomes. Mol Microbiol 60:1164–1178
- Fogel MA, Waldor MK (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. Genes Dev 20:3269–3282
- Garner EC, Campbell ChS, Mullins RD (2004) Dynamic instability in a DNA-segregating prokaryotic actin homolog. Science 306:1021–1025
- Ghosh S, Hajra S, Paek A, Jayaram M (2006) Mechanisms for chromosome and plasmid segregation. Annu Rev Biochem 75:211–241
- Gitai Z, Thanbichler M, Shapiro L (2005) The choreographed dynamics of bacterial chromosomes. Trends Microbiol 13:221–228
- Golding I, Cox EC (2004) RNA dynamics in live *Escherichia coli* cells. Proc Natl Acad Sci USA 101:11310–11315

- Goodsell DS (1993) The machinery of life. Springer-Verlag, New York
- Gordon GS, Sitnikov D, Webb ChD, Teleman A, Straight A, Losick R, Murray AW, Wright A (1997) Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* 90:1113–1121
- Gordon GS, Wright A (2000) DNA segregation in bacteria. *Annu Rev Microbiol* 54:681–708
- Guacci V (2007) Sister chromatid cohesion: the cohesion cleavage model does not ring true. *Genes Cells* 12:693–708
- Hansen FG, Christensen BB, Atlung T (1991) The initiator titration model: computer simulation of chromosome and minichromosome control. *Res Microbiol* 142:161–167
- Hecht RM, Stimpson D, Pettijohn D (1977) Sedimentation properties of the bacterial chromosome as an isolated nucleoid and as an unfolded DNA fiber. *J Mol Biol* 111:257–277
- Helmstetter CE, Cooper S (1968) DNA synthesis during the division cycle of rapidly growing *Escherichia coli* B/r. *J Mol Biol* 31:507–518
- Higgins NP (1999) DNA supercoiling and its consequences for chromosome structure and function. In: Charlebois RL (ed) Organization of the prokaryotic genome, Chapter 11. American Society for Microbiology, Washington, DC, pp 189–202
- Jensen RB, Shapiro L (1999) The *Caulobacter crescentus smc* gene is required for cell cycle progression and chromosome segregation. *Proc Natl Acad Sci USA* 96:10661–10666
- Jensen RB (2006) Coordination between chromosome replication, segregation and cell division in *Caulobacter crescentus*. *J Bacteriol* 188:2244–2253
- Jeong KS, Ahn J, Khodursky AB (2004) Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol* 5:R86
- Jin DJ, Cabrera JE (2006) Coupling the distribution of RNA polymerase to global gene regulation and the dynamic structure of the bacterial nucleoid in *Escherichia coli*. *J Struct Biol* 156:284–291
- Jun S, Mulder B (2006) Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proc Natl Acad Sci USA* 103:12388–12393
- Karczmarek A, Martinez-Arteaga Baselga R, Alexeeva S, Hansen FG, Vicente M, Nanninga N, den Blaauwen T (2007) DNA and origin region segregation are not affected by the transition from rod to sphere after inhibition of *Escherichia coli* MreB by A22. *Mol Microbiol* 65:51–63
- Kitamura E, Blow JJ, Tanaka TU (2006) Live-cell imaging reveals replication of individual replisomes in eukaryotic replication factories. *Cell* 125:1297–1308
- Kornberg T, Lockwood A, Worcel A (1974) Replication of the *Escherichia coli* chromosome with a soluble enzyme system. *Proc Natl Acad Sci USA* 71:3189–3193
- Li Y, Sergueev K, Austin S (2002) The segregation of the *Escherichia coli* origin and terminus of replication. *Mol Microbiol* 46:985–995
- Luijsterburg MS, Noom MC, Wuite GJL, Dame RT (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol* 156:262–272
- Maaløe O, Kjeldgaard NO (1966) The control of macromolecular synthesis. W.A. Benjamin, New York
- Marko JF, Siggia ED (1997) Polymer models of meiotic and mitotic chromosomes. *Mol Biol Cell* 8:2217–2231
- Mason DJ, Powelson DM (1956) Nuclear division as observed in live bacteria by a new technique. *J Bacteriol* 71:474–479
- Murphy LD, Zimmerman SB (2000) Multiple restraints to the unfolding of spermidine nucleoids from *Escherichia coli*. *J Struct Biol* 132:46–62
- Murphy LD, Zimmerman SB (2002) Hypothesis: the RNase-sensitive restraint to unfolding of spermidine nucleoids from *Escherichia coli* is composed of cotranslational insertion linkages. *Biophys Chem* 101:321–331
- Nanninga N (1969) Preservation of the ultrastructure of *Bacillus subtilis* by chemical fixation as verified by freeze-etching. *J Cell Biol* 42:733–744
- Nanninga N, Roos M, Woldringh CL (2002) Models on stickiness of replicated *Escherichia coli oriC*. *Microbiology* 148:3327–3328

- Nasmyth K (2002) Segregating sister genomes: the molecular biology of chromosome separation. *Science* 297:559–565
- Nielsen HJ, Li Y, Youngren B, Hansen FG, Austin SJ (2006a) Progressive segregation of the *Escherichia coli* chromosome. *Mol Microbiol* 61:383–393
- Nielsen HJ, Ottesen JR, Youngren B, Austin SJ, Hansen FG (2006b) The *Escherichia coli* chromosome is organized with the left and right chromosome arms in separate cell halves. *Mol Microbiol* 62:331–338
- Nielsen HJ, Li Y, Youngren B, Hansen FG, Austin SJ (2007) Dynamics of *Escherichia coli* chromosome segregation during multifork replication. *J Bacteriol* 189:8660–8666
- Niki H, Yamaichi Y, Hiraga S (2000) Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14:212–223
- Noom MC, Navarre WW, Oshima T, Wuite GJL, Dame RT (2007) H-NS promotes looped domain formation in the bacterial chromosome. *Curr Biol* 17:R913–R914
- Norris V, Madsen MS (1995) Autocatalytic gene expression occurs via transertion and membrane domain formation and underlies differentiation in bacteria: a model. *J Mol Biol* 253:739–748
- Odijk T (1998) Osmotic compaction of supercoiled DNA into a bacterial nucleoid. *Biophys Chem* 73:23–30
- Ohniwa RL, Morikawa K, Kim J, Kobori T, Hizume K, Matsumi R, Atomi H, Imanaka T, Ohta T, Yoshimura SH, Takeyasu K (2007) Atomic force microscopy dissects the hierarchy of genome architectures in eukaryote, prokaryote and chloroplast. *Microsc Microanal* 13:3–12
- Ortiz JO, Förster F, Kürner J, Linaroudis AA, Baumeister W (2006) Mapping 70S ribosomes in intact cells by cryoelectron tomography and pattern recognition. *J Struct Biol* 156:334–341
- Pettijohn DE (1996) The nucleoid. In: Neidhardt et al. FC (eds) *Escherichia coli* and *Salmonella*: cellular and molecular biology. American Society of Microbiology, Washington, DC, pp 158–166
- Poindexter JS (1964) Biological properties and classification of the *Caulobacter* group. *Bacteriol Rev* 28:231–295
- Reisner W, Morton KJ, Riehn R, Wang YM, Yu Z, Rosen M, Sturm JC, Chou SY, Frey E, Austin RH (2005) Statics and dynamics of single DNA molecules confined in nanochannels. *Phys Rev Lett* 94:196101-1-4
- Reyes-Lamothe R, Wang X, Sherratt D (2008a) *Escherichia coli* and its chromosome. *Trends Microbiol* 16:238–245
- Reyes-Lamothe R, Wang X, Sherratt D (2008b) Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* 133:90–102
- Riva A, Carpentier A-S, Barloy-Hubler F, Chéron A, Hénaut A (2008) Analyzing stochastic transcription to elucidate the nucleoid's organization. *BMC Genomics* 9:125
- Robinow C, Kellenberger E (1994) The bacterial nucleoid revisited. *Microbiol Rev* 58:211–232
- Romantsov T, Fishov I, Krichevsky O (2007) Internal structure and dynamics of isolated *Escherichia coli* nucleoids assessed by fluorescence correlation spectroscopy. *Biophys J* 92:2875–2884
- Roos M, van Geel ABM, Aarsman MEG, Veuskens JTM, Woldringh CL, Nanninga N (2001) The replicated *ftsQAZ* and *minB* chromosomal regions of *Escherichia coli* segregate on average in line with nucleoid movement. *Mol Microbiol* 39:633–640
- Shen V, Bremer H (1977) Chloramphenicol-induced changes in the synthesis of ribosomal, transfer, and messenger ribonucleic acids in *Escherichia coli* B/r. *J Bacteriol* 130:1098–1108
- Stavans J, Oppenheim A (2006) DNA-protein interactions and bacterial chromosome architecture. *Phys Biol* 3:R1–R10
- Toro E, Hong S-H, McAdams HH, Shapiro L (2008) *Caulobacter* requires a dedicated mechanism to initiate chromosome segregation. *Proc Natl Acad Sci USA* 105:15435–15440
- Trun NJ, Marko JF (1998) Architecture of a bacterial chromosome. *Am Soc Microbiol News* 68:276–283
- Valens M, Penaud S, Rossignol M, Cornet F, Boccard F (2004) Macrodome organization of the *Escherichia coli* chromosome. *EMBO J* 23:4330–4341

- Valkenburg JAC, Woldringh CL (1984) Phase separation between nucleoid and cytoplasm in *Escherichia coli* as defined by immersive refractometry. *J Bacteriol* 160:1151–1157
- van Helvoort JMLM, Huls PG, Vischer NOE, Woldringh CL (1998) Fused nucleoids resegment faster than cell elongation in *Escherichia coli* *pbpB*(Ts) filaments after release from chloramphenicol inhibition. *Microbiology* 144
- van Iterson W (1965) Symposium on the fine structure of bacteria and their parts. II. Bacterial cytoplasm. *Bacteriol Rev* 29:290–325
- Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Nat Acad Sci USA* 101:9257–9262
- Wang X, Possoz C, Sherratt DJ (2005) Dancing around the divisome: asymmetric chromosome segregation in *Escherichia coli*. *Genes Dev* 19:2367–2377
- Wang X, Liu X, Possoz C, Sherratt DJ (2006) The two *Escherichia coli* chromosome arms locate to separate cell halves. *Genes Dev* 20:1727–1731
- Wang X, Reyes-Lamothe R, Sherratt DJ (2008) Modulation of *Escherichia coli* sister chromosome cohesion by topoisomerase IV. *Genes Dev* 22:2426–2433
- Webb CD, Graumann PL, Kabana JA, Teleman AA, Silver PA, Losick R (1998) Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28:883–893
- White MA, Eykelenboom JK, Lopez-Vernaza MA, Wilson E, Leach DRF (2008) Non-random segregation of sister chromosomes in *Escherichia coli*. *Nature* 455:1248–1250
- Woldringh CL, de Jong MA, van den Berg W, Koppes L (1977) Morphological analysis of the division cycle of two *Escherichia coli* substrains during slow growth. *J Bacteriol* 131:270–279
- Woldringh CL, Nanninga N (1985) Structure of nucleoid and cytoplasm in the intact cell. In: Nanninga N (ed) *Molecular cytology of Escherichia coli*. Academic, London/New York, pp 161–197
- Woldringh CL, Zaritsky A, Grover NB (1994) Nucleoid partitioning and the division plane in *Escherichia coli*. *J Bacteriol* 176:6030–6038
- Woldringh CL, Odijk T (1999) Structure of DNA within the bacterial cell: physics and physiology. In: Charlebois RL (ed) *Organization of the prokaryotic genome*, Chapter 10. American Society for Microbiology, Washington, DC, pp 171–187
- Woldringh CL, van Driel R (1999) The eukaryotic perspective: similarities and distinctions between pro- and eukaryotes. In: Charlebois RL (ed) *Organization of the prokaryotic genome*, Chapter 5. American Society for Microbiology, Washington, DC, pp 77–90
- Woldringh CL (2002) The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. *Mol Microbiol* 45:17–29
- Woldringh CL, Nanninga N (2006) Structural and physical aspects of bacterial chromosome segregation. *J Struct Biol* 156:273–283
- Worcel A, Burgi E (1972) On the structure of the folded chromosome of *E. coli*. *J Mol Biol* 71:127–147
- Zaritsky A, Woldringh CL, Fishov I, Vischer NOE, Einav M (1999) Varying division planes of secondary constrictions in spheroidal *Escherichia coli* cells. *Microbiol* 145:1015–1022
- Zimmerman SB, Murphy LD (1996) Macromolecular crowding and the mandatory condensation of DNA in bacteria. *FEBS Lett* 390:245–248
- Zimmerman SB (2002) Toroidal nucleoids in *Escherichia coli* exposed to chloramphenicol. *J Struct Biol* 138:199–206
- Zimmerman SB (2006a) Cooperative transitions of isolated *Escherichia coli* nucleoids: implications for the nucleoid as a cellular phase. *J Struct Biol* 153:160–175
- Zimmerman SB (2006b) Shape and compaction of *Escherichia coli* nucleoids. *J Struct Biol* 156:255–261

Chapter 6

Polymer Physics for Understanding Bacterial Chromosomes

Suckjoon Jun

Abstract This chapter presents polymer physics that is relevant for an understanding of bacterial chromosome segregation. I first show that polymers have a natural tendency for segregation, which can be very strong in the presence of confinement. I then discuss segregation of duplicating chromosomes using the concentric-shell model, which predicts that newly synthesized DNA will be found in the periphery of the chromosome during replication. I sketch implications of these results, e.g., on the role of proteins, segregation mechanisms for bacteria of diverse shapes, cell cycle of an artificial cell, and evolution. Finally, I remind the reader of the robust nature of the bacterial cell cycle by presenting experimental results of *Escherichia coli* growing in a nano-fabricated slab of thickness under 300 nm.

Keywords Chromosome segregation • nucleoid compaction • polymer physics • molecular crowding

6.1 Introduction

Chromosome segregation is one of the fundamental and yet most elusive processes of the bacterial cell cycle. The model system *Escherichia coli* contains a single circular chromosome in a rod-shaped cell. DNA replication starts at a unique origin of replication (*oriC*), creating a replication bubble that grows bidirectionally, and the two replication forks meet at the terminus (*ter*) located at approximately the opposite clock position of *oriC* on the circular chromosome (Fig. 6.1). In slowly growing cells, there is a one-to-one correspondence between each complete round of replication and the cell cycle. In fast growing cells, on the other hand, the cell divides more frequently than the progression of the forks from *oriC* to *ter* and, thus, new replication has to initiate before the completion of the previous round of duplication, leading to “multifork” replication. In either case, replication and segregation are coordinated with the growth of the cell, and recent

S. Jun (✉)

FAS Center for Systems Biology, Harvard University, 52 Oxford St, Cambridge, MA, 02138, USA

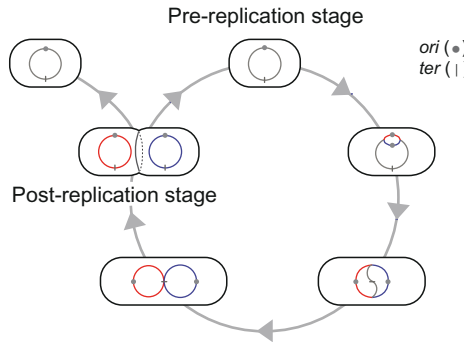


Fig. 6.1 Schematic illustration of the bacterial cell cycle. Replication begins at the 12' position (*ori*) on the circular chromosome. Replication forks grow bidirectionally and meet at the opposite clock position (*ter*). During the bacterial cell cycle, replication and segregation progress hand-in-hand, coupled with the growth of the cell

visualization experiments have begun to reveal how replicating chromosomes move and segregate during the bacterial cell cycle. Although details vary from organism to organism, the observation common to all bacteria of rod-shaped cells studied so far shows directed movement of duplicating chromosomes along the long axis of the cell as seen in *E. coli* (Bates and Kleckner 2005; Elmore et al. 2005; Nielsen et al. 2006; Wang et al. 2006), *Caulobacter crescentus* (Toro et al. 2008; Viollier et al. 2004), *Bacillus subtilis* (Berkmen and Grossman 2006; Berlatzky et al. 2008) and *Vibrio cholerae* (Fiebig et al. 2006; Fogel and Waldor 2006). Experimental data also suggest that, once duplicated, there is a fair degree of correlation between the relative clock position of circular chromosomes and their *average* positions along the long axis of the cell, which often (but not always) shows the principal linear ordering of chromosomes (Teleman et al. 1998; Viollier et al. 2004; Wang et al. 2006).

Perhaps the most influential model of chromosome segregation so far was proposed by Jacob et al. (1963) in their seminal paper on the replicon model of *E. coli*: if replicating chromosomes (i.e. duplicated *ori*'s) are attached to the elongating cell-wall membrane, they can be segregated passively by insertion of membrane material between the attachment points during growth. The model is intuitive and elegant – but wrong.

Since the work of Jacob and colleagues, two classes of models have been proposed to explain the observed directed drift of duplicating chromosomes: (a) biological mechanisms such as DNA replication (Lemon and Grossman 2001), co-transcriptional translation of membrane proteins (Woldringh 2002), RNA transcription (Dworkin and Losick 2002); and (b) physical/mechanical driving forces such as mechanical pushing between chromosomes (Bates and Kleckner 2005; Kleckner et al. 2004) or conformational entropy of duplicating chromosomes (Danchin et al. 2000; Fan et al. 2007; Jun and Mulder 2006). These two classes of model are not mutually exclusive.

The purpose of this chapter is to present polymer physics that is relevant for an understanding of bacterial chromosome segregation.¹ Based on the model I describe below, I posit that replicating *E. coli* daughter DNA strands, despite their strong confinement within the cell, will spontaneously demix as a result of entropic forces helped by proteins; i.e., entropy can act as a physical force that drives chromosome segregation.

6.2 Physical Basis for Bacterial Chromosomes in Confinement

6.2.1 Entropy Measures the Degrees of Freedom of the System

In his influential book, *What is life?* Schrödinger (1944) asserted that the only thermodynamically equilibrium state of a living being is death. Defining entropy as a measure of ‘disorder’, in the last chapter ‘Is Life Based on the Laws of Physics?’ he struggled to explain how order (life) can be achieved from disorder. This association of life with minimal entropy, however, can be misleading.

In Fig. 6.2a, I show snapshots from molecular dynamics simulations, where two species of N_1 and N_2 particles in equilibrium (blue and red, respectively) in a long rectangular box, initially separated by a wall. As we remove the wall from the system, the two species start to mix. The driving force of this process is the well-known “entropy of mixing,” which can be estimated as

$$\Delta S_A = k_B (\ln \Omega'_A - \ln \Omega_{A,0}) \quad (6.1)$$

$$\approx -k_B [N_1 \ln(\frac{N_1}{N}) + N_2 \ln(\frac{N_2}{N})] > 0, \quad (6.2)$$

where k_B is the Boltzman constant, $N = N_1 + N_2$, and $\Omega' = N!/N_1!N_2!$ ($\Omega_{A,0} = 1$) denotes the total number of configurations of the system after (before) mixing.

Entropy, however, is more subtle than a simple measure of disorder. To see this, let us start with the mixed state of Fig. 6.2a and connect the particles of the same species, creating two long linear chains, one painted with blue and the other with red. Another important condition is the excluded-volume interaction between the particles. The reader is encouraged to perform this simple computer simulation, and she or he will see that the two chains de-mix, i.e., that “order” emerges out of disorder.

Another example of order-out-of-disorder is the phase separation of a long polymeric chain (e.g., dsDNA) embedded in a solution of small molecules (e.g., globular

¹For a physical approach to eukaryotic chromosomes, the reader is encouraged to read Marko and Siggia (1997).

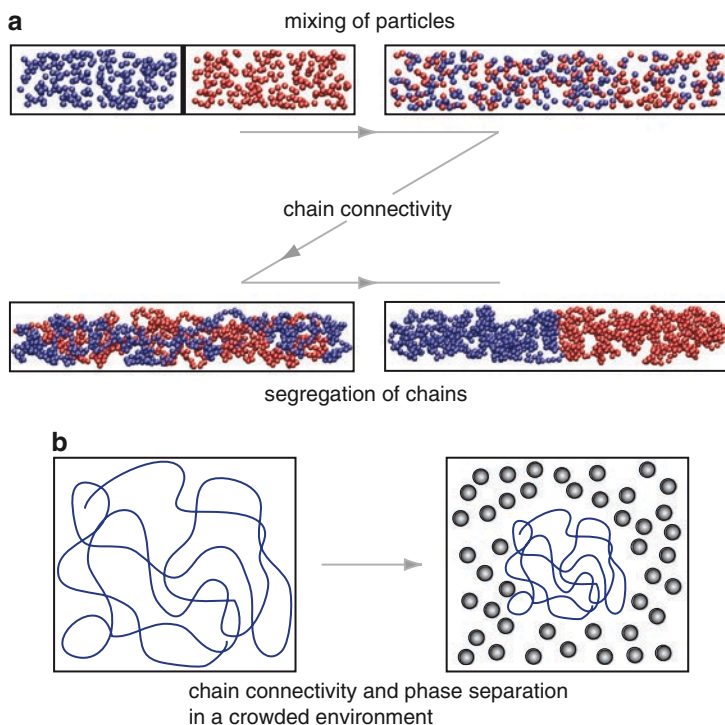


Fig. 6.2 An importance of chain connectivity (molecular dynamics simulation). **(a)** Two species of particles ($N_1 = N_2 = 200$), initially separated by a wall, mix as the wall is removed. If chain connectivity is introduced to this system by linearly connecting the particles of the same species, the two chains segregate ($N_1 = N_2 = 512$). **(b)** Crowding agents can collapse an imbedded self-avoiding chain to increase their translational entropy. The phase separation is not possible between two species of particles of the same size in the absence of chain connectivity

proteins), where the small molecules are depleted from the long chain because of their excluded volume (Fig. 6.2b). The result of this entropic interaction is the compaction of the long chain molecule (see Chapter 5).

Indeed, there is a plethora of examples in soft condensed matter physics where entropy leads to ordering. For instance, Onsager's hard-rod model (1949) of the nematic-isotropic transition of liquid crystals is based on a similar physical insight of the trade-off between loss of orientational entropy and gain in positional entropy when hard rods are oriented parallel to one another. Crystallization of hard spheres (Alder and Wainwright 1957; Wood and Jacobson 1957) is another example where ordering allows more room for fluctuations.

The above examples clarify what entropy really measures, namely, the degrees of freedom, or the size of the phase space, of the system. For the main theme of this article, the specific case in Fig. 6.2 points out the importance of chain connectivity, a keyword in polymer physics. In this example (Fig. 6.2b), chain connectivity defines

a conformational space Ω_B within the configuration space Ω_A , and within Ω_B the typical conformations of the chains are those de-mixed. This emergence of order from disorder due to chain connectivity and conformational entropy is our starting point of understanding chromosome segregation in bacteria.²

6.2.2 Polymers Have a Natural Tendency of Demixing, Which Can Be Stronger in Confinement

When two polymeric coils consisting of N monomers, each in dilute solution (in a good solvent), are brought together within a distance smaller than their natural size, both coils lose their conformational freedom because of the excluded-volume interactions and the chain connectivity and, thus, will resist overlap (Fig. 6.3a). This effective entropic repulsion between two chains was first characterized by Flory and Krigbaum in 1950 (Flory and Krigbaum 1950). The heart of their argument was to view the overlapping chains as a conceptual sphere of diameter of their radius of gyration $R_g \sim N^{3/5}$, which contains N monomers per chain. Their estimate of the free energy of collisions between the monomers was N^2 per volume, i.e., $F \sim N^2/R_g^3 \sim N^{1/5}$. Thus, the repulsion between the chains should grow as the length (N) of the chain grows, and each chain is an effective hard sphere. While their insight about polymeric repulsion was correct, it took another three decades for physicists to realize that the two chains in dilute bulk solution repel only weakly with order of $k_B T$ regardless of their sizes N (Grosberg et al. 1982; Jun et al. 2007). This is closely related to one of the basic ansätze in polymer physics that a single polymer blob,

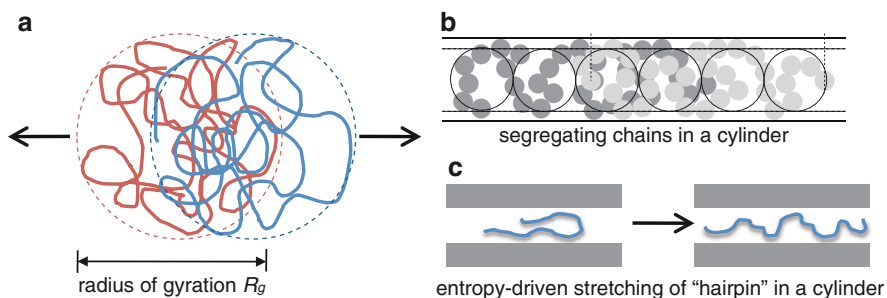


Fig. 6.3 Polymers have a natural tendency of repulsion. (a) In bulk, the free-energy barrier to overlap two blobs of chain is not high, i.e., repulsion is weak. (b) In a cylinder, a large chain “breaks” into smaller blobs, and the free energy barrier of chain mixing increases accordingly. (c) Experimentally, dsDNA with hairpin conformation has been shown to unfold driven by entropy (see text)

²Note that polymer physics was still in its infancy during Schrödinger’s time. For instance, Flory’s magnum opus, *Principles of Polymer Chemistry* (1953), appeared almost 10 years after *What is Life?* (1944).

regardless of its size, stores about $1k_B T$ of free energy. Note that $k_B T$ is the typical scale of thermal energy. For example, ATP hydrolysis releases about $\sim 12k_B T$.

In the presence of confinement, however, the effective repulsion can increase dramatically. To see this effect, consider two linear chains with excluded-volume interactions, which are initially intermingled and confined in an infinitely long cylinder of diameter $D \sim R_g$ (Fig. 6.3b). The free energy cost of intermingling in this case can be estimated by counting the total number of blobs in the tube (see above the “ $k_B T$ -per-blob” ansatz). Because the total number of blobs per chain is of order $D^{-5/3}N$, the overlapping free energy increases rapidly as the size of the chain increases. The repulsion between two chains in confinement is thus predicted to be very strong, and because of this entropic driving force the two chains slide away from each other at a characteristic rate $v \sim 1/DN$ over a distance comparable with the size of an unperturbed chain $l \sim D^{-2/3}N$, giving the timescale of segregation as $\tau_{\text{seg}} \sim l/v \sim N^2 D^{1/3}$. Note that this is much smaller (faster) than the typical timescale of random diffusion $\tau_{\text{diff}} \sim N^3 D^{-4/3}$ in a tube without entropic repulsion.

Similarly, in support of the above theoretical argument, recent experiments employing DNA molecules in nano-channels have demonstrated that dsDNA with hairpin conformation unfolds driven by its conformational entropy because the DNA tries to minimize the overlapping region (Fig. 6.3c) (Levy et al. 2009) (see also Odijk (2006) for other entropic effects). Future work will have to address quantitatively how the natural tendency of demixing of the chains will be influenced when the chains are in closed geometry as opposed to an open geometry such as a cylinder or a slab.

6.2.3 Polymers in a Thin Slab Segregate

So far, I have discussed polymer repulsion in bulk and in a cylinder. As I shall discuss later, another important geometry for chromosome segregation bacteria is a thin slab. In such two-dimensional confinement, polymers will segregate more strongly as the concentration of the chain increases (Fig. 6.4). This can be shown using the approach taken in Jun et al. (2007), which predicts that the free-energy cost of a chain overlapping crosses over from $\beta F \sim n^{d\nu_d/(d\nu_d-1)}$ to $\sim n^{d/(d-2)}$, where $\nu_d = 3/(d+2)$ is the Flory exponent and d the spatial dimension. thus, for $d = 3$,



Fig. 6.4 Chains in a two-dimensional slab segregate much more readily than in three dimensions

the free energy of chain mixing diverges at high chain concentration, i.e., chains will demix. Intuitively, segregation in 2D can be understood because individual chains cannot “percolate” from wall-to-wall without crossing other chains many times, which is costly. I will discuss two important implications of this result in the later sections.

6.2.4 *Effect of Chain Topology – Branched Polymers Repel Strongly*

The bacterial chromosome poses an interesting question concerning the effect of chain topology. First, the literal topology of most bacterial chromosomes is not linear but circular. In bulk solutions without strong confinement, it is known that there is an additional topological repulsion between ring polymers due to the non-concatenation condition (Cates and Deutsch 1986). In the presence of strong confinement, however, such “end effect” of the chain molecules will not contribute to the repulsive interaction between the chains significantly, since the correlation length of the chain becomes much smaller than the radius of gyration of the chain.

From a segregation point of view, the most important feature of the bacterial chromosome is the existence/presence of supercoiled plectonemes due to the under-twisting of the DNA-helix. This makes the effective topology of the bacterial chromosome mimic branched, tree-like structures. Branched polymers have higher internal topological complexity, which makes the repulsive interactions much stronger than between linear or circular chains. To see intuitively, imagine a polymer chain forming a square-lattice-like meshwork. Such objects are essentially 2-dimensional membranes, which cannot interpenetrate one another; whereas a simple linear chain can overlap only at the free energy barrier $\sim k_b T$ (Fig. 6.5). Branched chains stand between these two extreme cases.

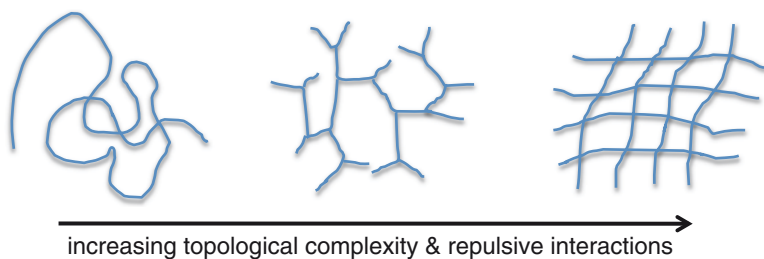


Fig. 6.5 Increasing internal topology of the chain corresponds to stronger repulsion between the chains

6.3 Chromosome Segregation in Bacteria

So far, I have discussed the physical conditions in which duplicated chromosomes will stay segregated, i.e., the conditions for a primordial driving force for chromosome segregation. How then can we explain the observed trajectories?

Figures 6.6 and 6.7 show experimental data of the average positions of the chromosome loci in *E. coli* (Bates and Kleckner 2005) and *C. crescentus* (Viollier et al. 2004). In *E. coli*, for instance, *oriC* first moves towards the midcell position and, then, replication starts. The duplicated *ori*'s split and move,³ on average, towards the 1/4 and 3/4 positions. In the meantime, *ter* drifts slowly from the cell pole towards the midcell, crossing one of the *ori* trajectories.

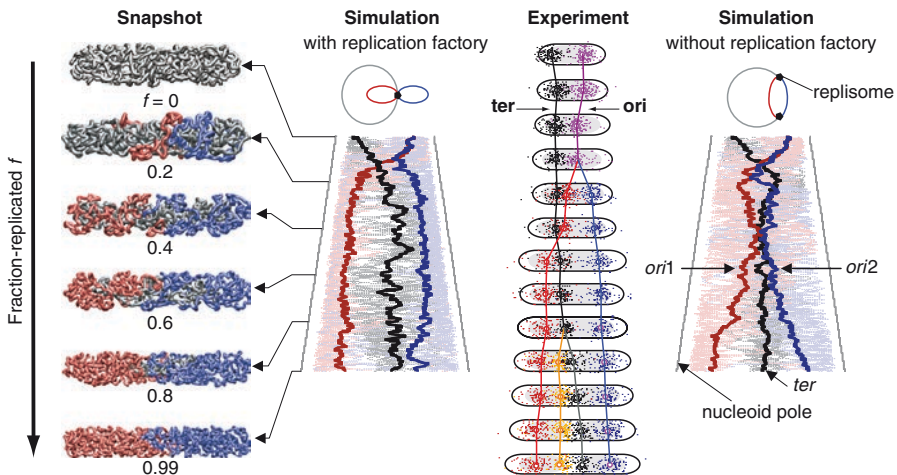


Fig. 6.6 Chromosome segregation in *E. coli*: simulation vs. experiment. (Left) A series of typical conformations of a replicating circular chain from 0% to 99% replication (un-replicated strand in gray, replicated strands in red and blue). We also present two sets of segregation pathways (*ori-ter* trajectories during replication) with and without “replication factory”; the third set of simulation where *ori* starts at the midcell can be found in Supporting Information in Jun et al. (2007). We simulated replication factories by enforcing physical proximity of the two replisomes during replication, but we did not fix their position within the cell. The dotted lines show the results of 10 individual simulations, and the thick solid lines show the average trajectories of *ori* (red and blue) and *ter* (black). (Center-to-Right) We juxtapose the simulations with the published data in Bates and Kleckner (2005) in an attempt to capture the main features of the experimental observations. For comparison, we used the fraction-replicated f as our “universal clock” and scaled the height of the simulation trajectories to match $0 < f < 1$ range of the data (Reprinted from Fig. 5 of Jun and Mulder (2006). Copyright (2006) National Academy of Sciences, U.S.A.)

³There is an ongoing debate about how long the duplicated *ori* and other loci stay together before splitting during chromosome segregation (“cohesion”). See, for example, Bates and Kleckner (2005), Nanninga et al. (2002), Nielsen et al. (2006), Sunako et al. (2001).

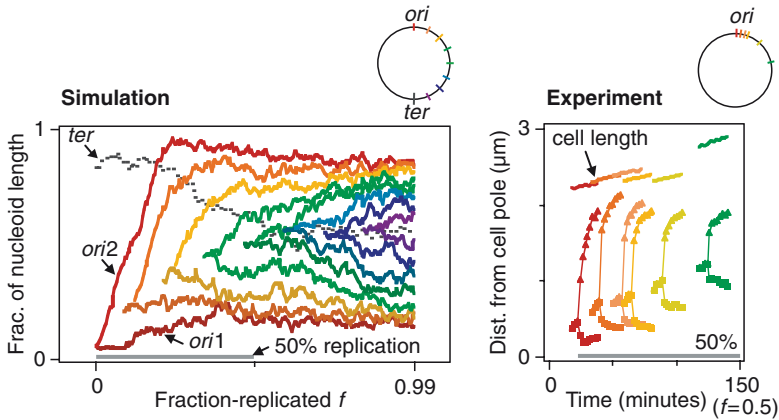


Fig. 6.7 Chromosome segregation of *C. crescentus*, comparing simulation (Left) vs. experiment Viollier et al. (2004) (Right). The simulated trajectories are the average of 26 individual simulation runs (or cells); we show the trajectories of nine representative loci (including *ori* and *ter*) on the right-arc of a circular chromosome for the entire duration of replication (up to 99.9%), whereas experimental data are only available for trajectories up to 50% of replication. For clarity, we show the trajectories only from the onset of replication of each locus. A full trajectory of *ter* is shown, however, to emphasize its slow drift from the cell pole to the cell center during replication, in contrast to the fast, directed diffusion of *ori2* in the nucleoid periphery. A main difference from the *E. coli* simulation is the additional assumption that we kept *ori1* in the volume near the stalked pole until 10–20% of the chain has been replicated. The concentric-shell model used in the simulation is formal and is not inconsistent with other additional mechanisms that may act on the directed movement of *ori2*, although in our simulation we did not need any such assumptions (Reprinted from Fig. 6 of Jun and Mulder (2006). Copyright (2006) National Academy of Sciences, U.S.A.)

In *C. crescentus*, loci trajectories seem even more striking (Viollier et al. 2004): one *ori* stays at the cell pole and the duplicated *ori* moves much faster than the growth rate of the cell (Fig. 6.7). Indeed, this large difference between the cell growth rate and the speed of *ori* is the evidence against Jacob and colleagues' model mentioned earlier, at least, in *C. crescentus*. The rest of the loci follow similar trajectories, creating a mirror-like image of the organization of duplicated circular chromosomes. *Ori* is found at the old cell pole and *ter* at the new cell pole; importantly, there is a one-to-one correspondence between the clock position of the chromosomal loci and their average physical locations along the long axis of the cell. In *E. coli*, recent data from several labs (e.g., Nielsen et al. 2006; Wang et al. 2006) also suggest similar principal linear organization, except that their clock positions are rotated by 90° with the two chromosome arms between *oriC* and *ter* occupying each cell half.

To explain these experimental data, we have previously proposed the concentric-shell model (Jun and Mulder 2006), which was inspired by the observation of nucleoid compaction in *E. coli* (see, also, Fig. 6.2b). Here, our prediction was that the newly synthesized DNA will move much faster in the periphery of the nucleoid near the cell-wall membrane than inside the nucleoid body (which is a meshwork

of chromosomal DNA), faster than the typical timescale of the cell cycle of *E. coli* (>20 min) or *C. crescentus* (≈ 240 min in Viollier et al. (2004)). Indeed, this minimal model and assumption reproduced most of the major features of the experimental data in both organisms.⁴ For *E. coli*, I note that Fan et al. (2007) have also used a free energy-driven string model to explain the experimental data. The most important result in their study is that the size of the chromosomal domain is important (the larger the domain is, the better the chromosomes segregate).

Note that the concentric-shell model is important not only to accomplish fast kinetics of DNA movement, but also to preserve the underlying ordering of the un-replicated mother strands. An elegant series of recent experiments, reported by Berlatzky et al. (2008), has confirmed our prediction. These workers labeled replicating chromosomes in slowly-growing living *B. subtilis* cells using fluorescent nucleotide derivatives, where DNA duplicated before (early) and after (late) a chosen timepoint was visualized by two colours. Although qualitative, their deconvolved images of early replicating DNA show its preferential distribution in the periphery of the nucleoid, resembling our simulation results (Fig. 6.6).

I emphasize that the concentric-shell model is consistent with other models, as long as they also imply the preferential occupation of the nucleoid periphery volume by newly synthesized DNA in the early stage of cell cycle. These models include, but are not limited to (1) the transertion model by Woldringh (2002), which assumes an interaction between the nucleoid and the plasma membrane via co-transcriptional translation and protein translocation, and (2) the putative role of *ParABS* on *ori* transport during the initial stage of the cell cycle in *Caulobacter* (Toro et al. 2008) and, potentially, in other organisms as well.

6.4 Discussion and Frequently Asked Questions

Although polymer physics offers important insights for the physical driving force underlying organization and segregation of bacterial chromosomes, living cells are more than a simple assembly of polymers in membrane. Below, I discuss the importance of proteins and various other questions concerning our polymer model of a bacterial chromosome.

6.4.1 Role of Proteins

Physical state of the chromosome. Various SMC-like or nucleoid-associated proteins such as MukBEF, HU, H-NS, IHF and gyrase may change the correlation length, i.e., the physical state of the chromosomes. For instance, nucleoid-associated

⁴In our previous study (Jun and Mulder 2006), this gap was even smaller than the width of the chain in our simulations.

proteins in general will increase the correlation length of the chromosome, because they will transfer the local density fluctuations of the DNA to longer distances. As discussed above, larger correlation lengths always result in more ordered conformation of the chain and better segregation. On the other hand, the more supercoiled the DNA is, the stronger the repulsion between the DNA strands becomes, due to the increased complexity of the chain topology. Thus, different microscopic and molecular details can help chromosome segregation when the physical conditions are met. From our point of view, the best example of this is in the work of Sawitzke and Austin (2000) (see also, Holmes and Cozzarelli (2000) and Dasgupta et al. (2000)). These authors have demonstrated that the severity of disruption of chromosome segregation due to Muk deficit is completely alleviated by alterations in gyrase activity leading to increased levels of negative supercoiling. Thus, neither class of proteins is the dedicated segregation machinery, but their role might be to help the chromosome stay in the right physical condition for segregation.

ParABS and kinetics and positioning of ori. There is increasing experimental evidence that *ori* and perhaps some other chromosomal loci are localized at specific intracellular positions in some organisms, e.g., *E. coli* (Danilova et al. 2007), *C. crescentus* (Toro et al. 2008; Viollier et al. 2004) and *V. cholerae* (Fiebig et al. 2006; Fogel and Waldor 2006). As we have shown previously using computer simulations (Jun and Mulder 2006), and also suggested by Breier and Cozzarelli (2004), pinning *ori* can help entropy set the orientation of the chromosome. This will be particularly pertinent for bacteria with more than one chromosome. Further, for the special case of cells with not enough space in the nucleoid periphery for fast diffusion of DNA, active *ori* transport and entropy may collaborate (see below). On the other hand, I note that there are special cases where proteins are involved in transport of DNA from one position to another inside a bacterial cell. Examples include SpoIIIE for sporulation in *B. subtilis* and FtsK for dimer resolution and other rescue tasks in *E. coli* (Barre 2007 and references therein). However, it is important to realize that these proteins translocate DNA, i.e., they take advantage of the directionality provided by the septum, which is entirely different from segregation of replicating chromosomes.

Cell growth, shape and division. Since the size and the aspect ratio of the cell are important (the higher the aspect ratio is, the faster chromosomes segregate), proteins that regulate cell growth and division are also important to ensure that the cell reaches the appropriate size for proper segregation. In a series of papers, Levin's lab has shown that *B. subtilis* possesses a metabolic sensor that couples nutritional availability to division, which serves to maintain a constant FtsZ ring to cell length ratio regardless of growth rate to ensure that cells reach the appropriate mass and complete chromosome segregation prior to cytokinesis (Weart et al. 2007). Similarly, other proteins can help segregation with their secondary effects, such as by measuring the length of the cell (e.g., MinCDE) or by maintaining a high aspect ratio of the cell (e.g., MreB). In addition to its role in maintaining a high aspect ratio of the cell,

MreB is believed to function in the final stages of chromosome segregation by regulating the activity of Topoisomerase IV at midcell (Madabhushi and Mariani 2009).

Molecular crowding and nucleoid compaction. During exponential phase, *E. coli* contains over one million proteins, occupying 10–20% of the cellular volume. These proteins, be they active or not, possess excluded volume and it has been predicted that their depletion interactions with the DNA results in the observed compaction of the nucleoid (Fig. 6.2b) (Woldringh and Odijk 1999). As I have discussed above, this molecular-crowding-induced phase separation between the nucleoid and the cytoplasm enables fast kinetics of newly synthesized DNA.

6.4.2 *Shouldn't a Hypothetical Motor Protein Be Enough to Segregate Chromosomes in Bacteria?*

No. There is increasing experimental evidence that *ori* and perhaps some other chromosomal loci are localized or tethered at specific intracellular positions (see, for example, Stavans and Oppenheim (2006) and references therein). However, one still must explain how the hypothetical transportation of a small fraction (e.g., *ori*) of millions of basepairs of DNA from one position inside the cell to another can dictate, if any at all, the directed movement, segregation and organization of the rest of the chromosome. These observations thus require understanding of more basic physical principles.

I note that there are special cases where motor proteins are indeed involved in transportation of DNA from one position to another inside a bacterial cell. The examples include *SpoIIIE* for sporulation in *B. subtilis* and *FtsK* for dimerresolution and other “rescue” tasks in *E. coli* (Barre 2007 and references therein). However, it is important to realize that these proteins *translocate* DNA, i.e., they take advantage of the directionality provided by the septum, which is entirely different from segregation of replicating chromosomes.

6.4.3 *How About Topologically Catenated Chromosomes?*

Imagine two topologically concatenated ring polymers confined in a rectangular box. These two ring polymers will still segregate to occupy each half of the box, i.e., the topology of the system localizes on average at the centre of the box (unpublished results). It is an intriguing question whether this localization of chain topology, which is a *global* property of the system, will influence the action of topoisomerases, which have access to only *local* information. I propose that the entropy-driven directionality of the movements and relative orientation of the confined

chains will, at least in part, help topoisomerases decatenate the replicated chains before cell division.

6.4.4 Bacterial Cells Are Very Crowded: Will Molecular Crowding Influence Chromosome Organization and Segregation?

In the first approximation, no. One may foresee two opposing arguments: (A) As the chains are compressed by depletion effect, the inner space occupied by the chains has a higher density of polymers, i.e., less accessible volume by the other chain and, thus, the tendency of de-mixing will increase because of squeezing, or (B) Molecular crowding exerts an effective osmotic pressure, which will make the chains mix.

In fact, a simple calculation suggests that there is a balance between these two effects. To see this, let us consider the free energy of the compressed chains in Fig. 6.8 as follows (Cacciuto and Luitjen 2006; Grosberg and Khokhlov 1994; Jun et al. 2007; Sakaue and Raphaël 2006).

$$\beta F_A = 2C \left(\frac{Nv}{D} \right)^{\frac{3}{3v-1}} \quad (6.3)$$

$$\beta F_B = C \left[\frac{(2N)v}{D'} \right]^{\frac{3}{3v-1}}, \quad (6.4)$$

where C is the prefactor. These two free energies for the compressed chains become equal at $D' = 2^{1/3}D$, i.e., when their total volumes are the same, $V_A = V_B$.

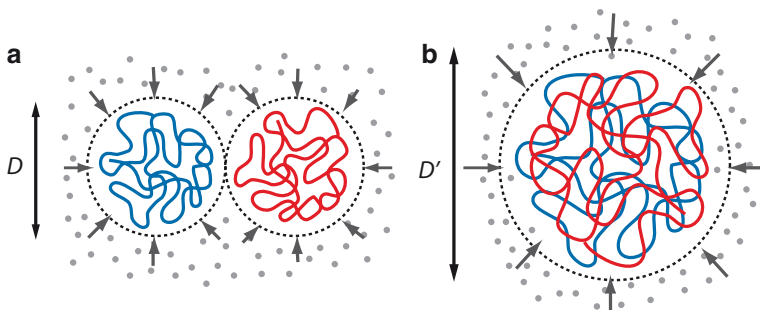


Fig. 6.8 Two possible scenarios of the role of molecular crowding on chain interactions. (a) Segregation (b) mixing. Each chain consists of N monomers. Calculations show that crowding, in the first approximation, should not influence mixing vs. de-mixing of the chains (see text)

Thus, compression due to depletion by itself will not influence mixing vs. de-mixing of the chains, even if it could bring the chains together and change their envelope shape (surface tension).⁵ Computer simulation results also support the neutral effect of molecular crowding (Axel Arnold, personal communication).

On the other hand, crowding can influence the local organization of the chromosome. For instance, looping is one way to achieve compaction of chromosomes, and the entropy gain by depletion attraction between DNA segments can be larger than the entropy loss by DNA looping (Marenduzzo et al. 2006).

6.4.5 *Bacteria Exist in Various Cell Shapes and Composition of Chromosomes. Are There General Strategies for Successful Chromosome Segregation in Bacteria?*

Since chromosome segregation is one of the defining processes of any cell, its basic mechanism must have been conserved across branches of life and organisms of diverse shapes. Below are some speculations.

Filamentous bacteria. I propose an entropy-driven, “random” segregation mechanism for filamentous bacteria. As I have shown both analytically and numerically, polymers confined in a narrow cylindrical geometry strongly resist overlap and, thus, repel one another (Arnold and Jun 2007; Jun and Mulder 2006), where the timescale of disentanglement of overlapping polymers is much shorter than that of pure diffusion. Indeed, recent experimental study on the filamentous cyanobacterium *Anabaena sp. PCC 7120* has revealed Gaussian distribution of DNA content in each daughter cells after the septum formation (Hu et al. 2007). Importantly, the variation of the distribution was much larger than a value expected if the two daughter cells were identical, suggesting a random event involved in chromosome segregation. Also, *Anabaena sp. PCC 7120*, like many other filamentous bacteria, is polyploid and contains multiple copies of chromosomes per cell, ~10 (Hu et al. 2007), which seems to moderate the effect of random segregation. Since chromosomes occupy a much larger volume than plasmids inside the cell, this apparent random segregation process is sufficient for the viability of the cell as long as multiple copies of chromosomes are produced before division.

Rod-shaped bacteria. Based on our previous work on *E. coli* and other work to be published elsewhere, we have critically examined the *E. coli* B/r strain and explained why duplicated chromosomes will not mix. Thus, other organisms of similar cell shape and

⁵The origin of this result is due to the form of the free energy in Eq. 6.3 (Grosberg et al. 1982; Jun et al. 2007; Sakaue and Raphaël 2006), which has been constructed to be a function of only the monomer density. In other words, regardless of the shape of the envelope surrounding the chains, the free energies will be the same as long as the volumes are also the same, and vice versa. Additional consideration of the interactions at the surface will change the envelope shape, but not segregation/mixing.

volume must benefit from the entropy-driven driving force of segregation, with or without any organism-specific segregation mechanisms, as long as the correlation of chromosome is also comparable to that of *E. coli* (≈ 100 nm).

Square bacteria. Perhaps one of the most striking bacterial shapes is that of Walsby's square bacterium (Walsby 2000). This organism has a very thin and square-shaped cell (≈ 100 nm thick and ≈ 2 μ m wide with aspect ratio 1), a "2-dimensional" creature resembling a postage stamp. Interestingly, microscopy images so far suggest that that square bacteria grow and divide at alternating perpendicular axes and planes, respectively, and the cells are typically observed in 2×2 , 4×4 , 8×8 stamp-like configurations. As I have explained above, polymers in a thin slab segregate strongly. Thus, simple symmetry breaking by cell growth and the thin 2D geometry may suffice to separate the chromosomes entropically.

Spherical bacteria. The perfect symmetry of the cell shape means that the confined chains do not have any preferred conformations between mixing and demixing, although their global reorganization could readily be achieved (Jun et al. 2007). Since few data are available on chromosome organization in spherical bacteria, I can only speculate about possible contributing factors to segregation. At the chromosome level, supercoiling-induced branched structures of the chromosome will increase the tendency of demixing (Jun and Mulder 2006; Trun and Marko 1998; Vilgis 2000). Further, symmetry breaking of the cell shape and invagination of the cell during division could help resolve partially intermingled chromosomes. Indeed, real cells are never perfectly spherical, and Huang et al. (2004) have shown numerically that Min-protein oscillations can be achieved along the long axis of the *nearly* round cell, where the equatorial radii differ by as little as 5%. This may explain the observed division of round cells at alternating perpendicular planes (Corbin et al. 2002). Along with the case of the filamentous bacteria discussed above, it is tempting to speculate that chromosome segregation in spherical bacteria is also a "random" process and, thus, polyploidy can be a strategy to increase the chance of successful distribution of the chromosomes to daughter cells. For example, the endosymbiotic bacteria of aphids, *Buchnera*, which cannot divide by itself outside the host eukaryotic cell, contains over 100 copies of its genome (Komaki and Ishikawa 1999).

6.4.6 Early Life and Artificial Cells

Previously, we proposed that the entropy-driven segregation of polymers could be implemented in designing an artificial cell, and that it has implications on early life (Jun and Mulder 2006). How could an artificial cell achieve its basic cellular processes? While Szostak and colleagues have proposed a system of a self-replicating vesicle and replicases as a protocell (Szostak et al. 2001), the cell division mechanism as well as the size regulation of daughter cells of a protocell are far from being understood.

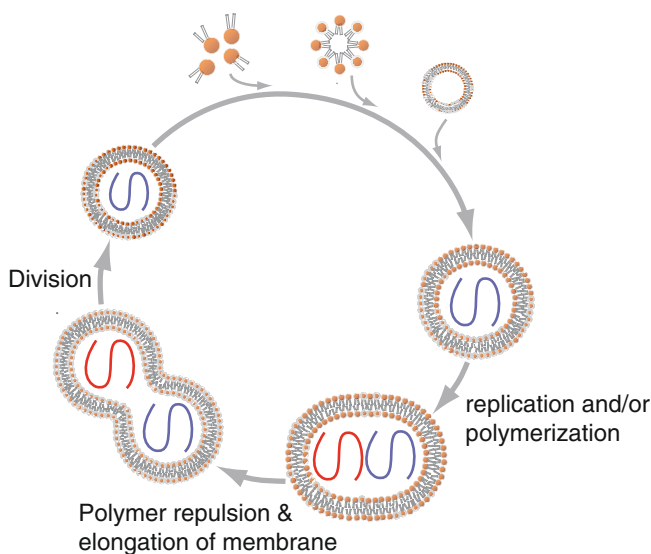


Fig. 6.9 Entropy-driven segregation of encapsulated polymers and its potential role on the cell cycle of a protocell. Repulsion between the replicated chains can break the spherical symmetry of the vesicle, and, thus, may facilitate its spontaneous division. Moreover, the daughter vesicles containing the duplicated polymers may have similar sizes. (The three minisketches at the top represent input from the environment for feeding the growth.) These scenarios could be tested both theoretically and experimentally (adapted from Figure 3 in Szostak et al. (2001))

I suggest that two long polymers, formed by replication/ligation of smaller molecules within a spherical vesicle repel each other because of entropy. If the polymer-polymer repulsion is strong enough to break the spherical symmetry of the vesicle, this process may lead to membrane fission, preferentially at the mid-cell position, which is defined by the surface of contact between the two repelling polymers. This would further help regulate the size distribution of newly formed daughter cell as illustrated in Fig. 6.9 (see also Chen 2009).

6.5 Concluding Remarks on the Robustness of Bacterial Cell Cycle

Finally, I would like to remind the reader that segregation by entropic forces represents a robust mechanism as we have learned from our recent experiments. The aim of these experiments was to observe growth and division of *E. coli* for many generations in a patterned surface. In the early stage of the development, we did not realize that there was a significant friction between the patterned surface and the cells, which caused abnormal growth of the cells. This, however, revealed an unexpected ability of *E. coli* to grow under extreme spatial constraints. Indeed, *E. coli* can go

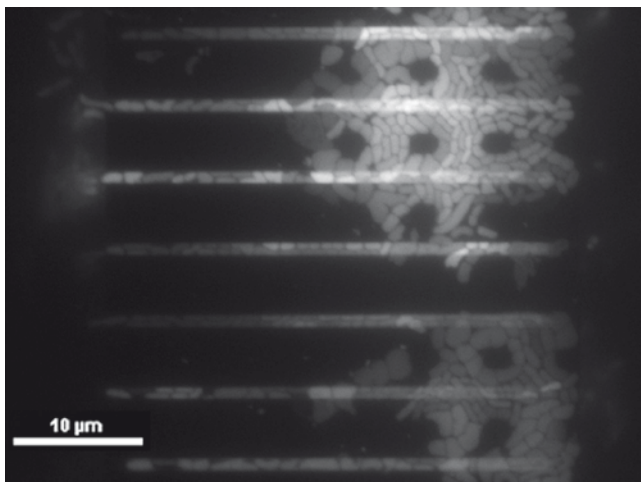


Fig. 6.10 Robustness of bacterial cell cycle. *E. coli* growing in extreme spatial constraints. The cells in the microchannels are often indefinable. When the channels are packed, the cells start to grow into narrow slits of width much smaller than the diameter of the normal *E. coli* cells. These cells have “pancake-like” morphology ($\approx 5\ \mu\text{m}$ wide, and 200–300 nm thick) and they can still grow and divide for several generations (see Appendix). The pancake *E. coli*, after they grow back into the channels, appear to restore their rod-shaped morphology after several divisions

through several generations in diverse shapes ranging from a “potato” (shape that cannot be characterized by simple geometry) to a “pancake” ($\approx 300\ \text{nm}$ thin and $5\text{--}10\ \mu\text{m}$ wide) (Fig. 6.10), which reminds us of Walsby’s intriguing square bacterium discussed in the text. I wonder whether *E. coli* would have been able to survive without the robust, physical mechanisms underlying chromosome segregation.

6.6 Materials and Methods for Fig. 6.10

Microfluidic device. I used a standard lithography technique and facility at Cornell NanoScale Science & Technology Facility to micro-pattern the silicon surface with three different length scales. In Fig. 6.11, the region where cells grow consists of two layers of depths; the growth channels (region A) have a depth ranging between 0.8 and $1.2\ \mu\text{m}$ and width 1.5 and $2.5\ \mu\text{m}$. The closing ceiling (glass coverslip) is supported by nano posts (region B), which have height 200 and 300 nm. This nano slit (region C) between the patterned surface and the coverslip is large enough so that nutrient can continuously flow into the growth channels, but narrow enough so that the cells in principle cannot escape. Imaging was done through the coverslip side

Strain and growth. I used *E. coli* MG1655 strain containing a GFP expression vector (pGFPgcn4) with ampicillin resistance (Ito et al. Biochem. Biophys. Res.

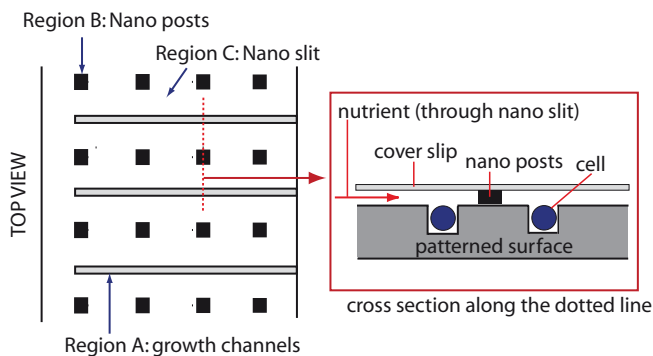


Fig. 6.11 Schematic drawing of the patterned surface and the position of growing of cells

Comm. 264, 556–560 (1999)). Cells were loaded to the micro-patterned surface (Fig. 6.11) and they were grown at 30°C. Growing cells were continuously supplied with Luria Broth (LB) pumped into the patterned surface (Fig. 6.11). Cells initially grew in the growth channels (region A). Due to significant friction between the cells and the surface of the channels walls, cells started to deform during growth. When the growth channels were completely packed, the cells start to grow into the nano-slit region (region C). Cells in region C thus have thickness less than 300 nm with several μm diameter, and most of them continued to grow and divide.

Acknowledgments I thank A. Arnold, B.-Y. Ha and B. Mulder for many of the collaborative results presented in this book chapter, and W. Dang and P. Galajda for their work mentioned in Fig. 6.10 during their stay in my lab. I am particularly grateful to C. Woldringh for numerous discussions, inspiring conversations and long-term collaborations.

References

- Alder BJ, Wainwright TE (1957) Phase transition for a hard sphere system. *J Chem Phys* 27:1208–1209
- Arnold A, Jun S (2007) Timescale of entropic segregation of flexible polymers in confinement: implications for chromosome segregation in filamentous bacteria. *Phys Rev E* 76:031901
- Barre FX (2007) FtsK and SpoIIIE: the tale of the conserved tails. *Mol Microbiol* 66:1051–1055
- Bates D, Kleckner N (2005) Chromosome and replisome dynamics in *E. coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. *Cell* 121:899–911
- Berkmen MB, Grossman AD (2006) Spatial and temporal organization of the *Bacillus subtilis* replication cycle. *Mol Microbiol* 62:57–71
- Berlatzky IA, Rouvinski A, Ben-Yehuda S (2008) Spatial organization of a replicating bacterial chromosome. *Proc Natl Acad Sci USA* 105:14136–14140
- Breier AM, Cozzarelli NR (2004) Linear ordering and dynamic segregation of the bacterial chromosome. *Proc Natl Acad Sci USA* 101:9175–9176

- Cacciuto A, Luijten E (2006) Self-avoiding flexible polymers under spherical confinement. *Nano Lett* 6:901–905
- Cates ME, Deutsch JM (1986) Conjectures on the statistics of ring polymers. *J Physique* 47:2121–2128
- Chen I (2009) Cell division: breaking up is easy to do. *Curr Biol* 19(8):R327–R328
- Corbin BD, Yu XC, Margolin W (2002) Exploring intracellular space: function of the Min system in round-shaped *Escherichia coli*. *EMBO J* 21:1998–2008
- Danchin A, Guerdoux-Jamet P, Moszer I, Nitschké P (2000) Mapping the bacterial cell architecture into the chromosome. *Phil Trans Roy Soc Lond B* 355:179–190
- Danilova O, Reyes-Lamothe R, Pinskaya M, Sherratt D, Possoz C (2007) MukB colocalizes with the *oriC* region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol Microbiol* 65:1485–1492
- Dasgupta S, Maisnier-Patin S, Nordström K (2000) New genes with old *modus operandi*: the connection between supercoiling and partitioning of DNA in *Escherichia coli*. *EMBO Reports* 1:323–327
- Dworkin J, Losick R (2002) Does RNA polymerase help drive chromosome segregation in bacteria? *Proc Natl Acad Sci USA* 99:14089–14094
- Elmore S, Müller M, Vischer N, Odijk T, Woldringh CL (2005) Single-particle tracking of *oriC*-GFP fluorescent spots during chromosome segregation in *Escherichia coli*. *J Struct Biol* 151:275–287
- Fan J, Tuncay K, Ortoleva PJ (2007) Chromosome segregation in *Escherichia coli* division: a free energy-driven string model. *Comp Biol Chem* 31:257–264
- Fiebig A, Keren K, Theriot JA (2006) Fine-scale time-lapse analysis of the biphasic, dynamic behaviour of the two *Vibrio cholerae* chromosomes. *Mol Microbiol* 60:1164–1178
- Flory PJ (1953) Principles of polymer chemistry. Cornell University Press, Ithaca, NY, USA
- Flory PJ, Krigbaum WR (1950) Statistical mechanics of dilute polymer solutions II. *J Chem Phys* 18:1086–1094
- Fogel MA, Waldor MK (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev* 20:3269–3282
- Grosberg AY, Khalatur PG, Khokhlov AR (1982) Polymeric coils with excluded volume in dilute solution: the invalidity of the model of impenetrable spheres and the influence of excluded volume on the rates of diffusion-controlled intermacromolecular reactions. *Die Makromol Chem, Rapid Commun* 3:709–713
- Grosberg AY, Khokhlov AR (1994) Statistical physics of macromolecules. AIP, New York, NY, USA
- Holmes VF, Cozzarelli NR (2000) Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc Natl Acad Sci USA* 97:1322–1324
- Hu B, Yang G, Zhao W, Zhang Y, Zhao J (2007) MreB is important for cell shape but not for chromosome segregation of the filamentous cyanobacterium *Anabaena sp. PCC 7120*. *Mol Microbiol* 63:1640–1652
- Huang KC, Wingreen NS (2004) Min-protein oscillations in round bacteria. *Phys Biol* 1:229–235
- Jacob F, Brenner S, Cuzin F (1963) On the regulation of DNA R replication in bacteria. *J Mol Biol* 28:329–348
- Jun S, Arnold A, Ha B-Y (2007) Confined space and effective interactions of multiple self-avoiding chains. *Phys Rev Lett* 98:128303
- Jun S, Mulder B (2006) Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proc Natl Acad Sci USA* 103:12388–12393
- Kleckner N, Zickler D, Jones GH, Dekker J, Padmore R, Henle J, Hutchinson J (2004) A mechanical basis for chromosome function. *Proc Natl Acad Sci USA* 101:12592–12597
- Komaki K, Ishikawa H (1999) Intracellular bacterial symbionts of aphids possess many genomic copies per bacterium. *J Mol Evol* 48:717–722
- Lemon KP, Grossman AD (2001) The extrusion-capture model for chromosome partitioning in bacteria. *Genes Dev* 15:2031–2041

- Levy SL, Mannion JT, Cheng J, Reccius C, Craighead HG (2009) Entropic unfolding of DNA molecules in nanofluidic channels. *Nano Lett* 8:3839–3844
- Madabhushi R, Mariani KJ (2009) Actin homolog MreB affects chromosome segregation by regulating topoisomerase IV in *Escherichia coli*. *Mol Cell* 33:171–180
- Marenduzzo D, Micheletti C, Cook PR (2006) Entropy-driven genome organization. *Biophys J* 90:3712–3721
- Marko JF, Siggia ED (1997) Polymer models of meiotic and mitotic chromosomes. *Mol Biol Cell* 8:2217–2231
- Nanninga N, Roos M, Woldringh CL (2002) Models on stickiness of replicated *Escherichia coli oriC*. *Microbiology* 148:3327–3328
- Nielsen HJ, Ottesen JR, Youngren B, Austin SJ, Hansen FG (2006) The *Escherichia coli* chromosome is organized with the left and right chromosome arms in separate cell halves. *Mol Microbiol* 62:331–338
- Odijk T (2006) DNA coned in nanochannels: Hairpin tightening by entropic depletion. *J Chem Phys* 125:204904
- Onsager L (1949) The effect of shape on the interaction of colloidal particles. *Ann NY Acad Sci* 51:627–659
- Sakaue T, Raphaël (2006) Polymer chains in confined spaces and flow-injection problems: some remarks. *Macromolecules* 39:2621–2628
- Sawitzke JA, Austin S (2000) Suppression of chromosome segregation defects of *Escherichia coli muk* mutants by mutations in topoisomerase I. *Proc Natl Acad Sci USA* 97:1671–1676
- Schrödinger (1944) What is life? Cambridge University Press, Cambridge, UK
- Stavans J, Oppenheim A (2006) DNA – protein interactions and bacterial chromosome architecture. *Phys Biol* 3:R1–R10
- Sunako Y, Onogi T, Hiraga S (2001) Sister chromosome cohesion of *Escherichia coli*. *Mol Microbiol* 42:1233–1241
- Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. *Nature* 409:387–390
- Teleman AA, Graumann PL, Lin DC-H, Grossman AD, Losick R (1998) Chromosome arrangement within a bacterium. *Curr Biol* 8:1102–1109
- Toro E, Hong S-H, McAdams HH, Shapiro L (2008) *Caulobacter* requires a dedicated mechanism to initiate chromosome segregation. *Proc Natl Acad Sci USA* 105:15435–15440
- Trun NJ, Marko JF (1998) Architecture of a bacterial chromosome. *ASM News* 64:276–283
- Vilgis TA (2000) Polymer theory: path integrals and scaling. *Phys Rep* 336:167–254
- Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Natl Acad Sci USA* 101:9257–9262
- Walsby AE (2000) A square bacterium. *Nature* 283:69–71
- Wang X, Liu X, Possoz C, Sherratt DJ (2006) The two *Escherichia coli* chromosome arms locate to separate cell halves. *Genes Dev* 20:1727–1731
- Weart RB, Lee AH, Chien A-C, Haeusser DP, Hill NS, Levin PA (2007) A metabolic sensor governing cell size in bacteria. *Cell* 130:335–347
- Woldringh CL (2002) The role of co-transcriptional translation and protein translocation (transer-tion) in bacterial chromosome segregation. *Mol Microbiol* 45:17–29
- Woldringh CL, Odijk T (1999) In: Charlebois (ed) Organization of the prokaryotic genome. ASM, Washington DC, USA, pp 161–197
- Wood WW, Jacobson JD (1957) Preliminary results from a recalculation of the Monte Carlo equation of state of hard spheres. *J Chem Phys* 27:1207–1208

Chapter 7

Molecular Structure and Dynamics of Bacterial Nucleoids

N. Patrick Higgins, B.M. Booker, and Dipankar Manna

Abstract This chapter summarizes recent evidence about the composition, structures, and function of components that form supercoil domains in bacterial nucleoids. Nucleoid formation enables a cell to move a compact DNA ensemble of dynamic transcription complexes and multiple replisomes efficiently into daughter cells during the cell division. In the past 10 years advances were made in many bacterial systems, but this review focuses on work carried out in the closely related clades of *Escherichia coli* and *Salmonella typhimurium*. New information indicates most of the chromosome has a stochastic domain structure but that transcription organizes unique topological zones in a small fraction of the genome. Nucleoid formation becomes critical when cells constrict the cell cycle to carry out dichotomous replication. Under very short doubling times of dichotomous growth, cell division requires intricate coordination of multiple factors to promote chromosome segregation. The most crucial enzyme for nucleoid formation is DNA gyrase, which not only removes positive links during DNA replication, it also establishes the supercoil tension that promotes correct compaction. Important enzymes including Topo IV, FtsK, and XerCD recombinases are tuned to promote nucleoid segregation only when appropriate levels of supercoiling are present. Surprisingly, gyrase establishes significantly different supercoiling levels in *E. coli* and *Salmonella typhimurium*, which makes the enzyme a barrier for horizontal gene transfer between the species.

Keywords Condensins • gyrase • supercoiling • supercoil diffusion • Topo I • Topo IV • domains • nucleoid • site-specific recombination • horizontal gene transfer • domainins • MukBEF • FtsK • replichores • *Salmonella* • *E. coli* • nucleoid-associated proteins • transcription • chromosome dynamics • DNA looping

N.P. Higgins (✉), B.M. Booker, and D. Manna
Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham,
Birmingham, AL 35226
e-mail: nphiggin@uab.edu

D. Manna
MicroPhage, Inc, 2400 Trade Center Avenue, Longmont, CO 80503

7.1 Introduction

What is negative supercoiling good for? Negative DNA supercoiling (–SC) generated by DNA gyrase contributes to physiological form and function of bacterial chromosomes. The first important role of –SC was discovered in biochemical studies of phage λ integration. λ site-specific recombination into the chromosomal bacterial attachment site (*attB*) required a supercoiled phage attachment site (*attP*). This discovery predicted an enzyme that can supercoil DNA and led to the discovery of gyrase (Gellert et al. 1976). Supercoiling was later found to be necessary for phage Mu transposition reactions (Craigie and Mizuuchi 1985), and today we know many site-specific recombination systems that rely on interwound supercoiling to form specific synaptic complexes that initiate site-specific DNA recombination reactions (Fig. 7.1).

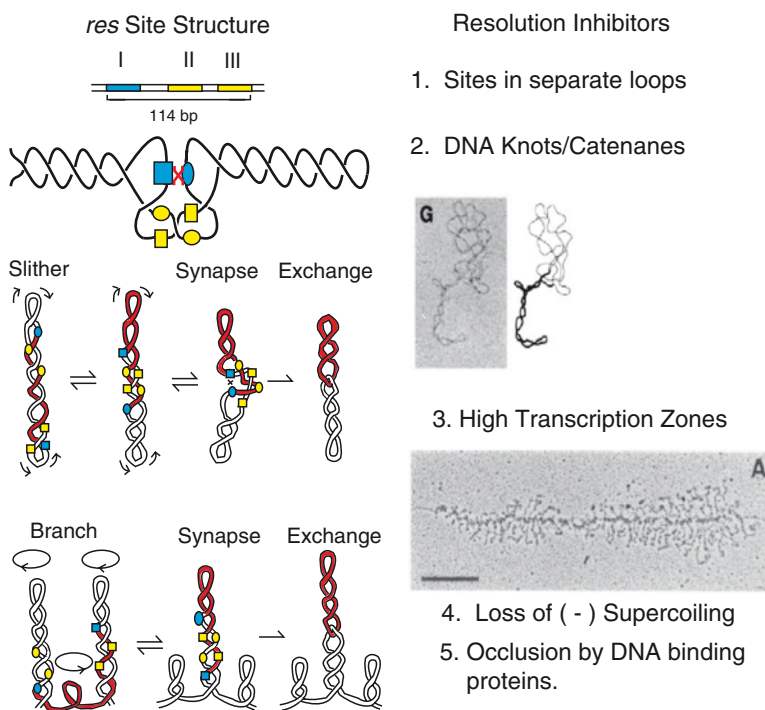


Fig. 7.1 The $\gamma\delta$ resolvase mechanism (left) and conditions that block synaptic pairing (right). A 114 bp $\gamma\delta$ *res* site includes 3 sub-sites labeled I (blue), II, and III (yellow). Each sub-site binds a resolvase dimer, shown as ovals or boxes on different directly repeated sites. Recombination requires a highly supercoiled synapse with 3 negative crossing nodes. Enzymes at *res* I (blue) can recombine while the other two sites (II and III) stabilize the complex during strand transfer. Two DNA movements bring sites into synapse: slithering (middle) or branching (bottom). Conditions that block synaptic pairing for recombination include: (1) sites separated in different loop domains. (2) Pre-catenane tangling of strands during DNA replication (Peter et al. 1998). (3) Genes undergoing high transcription (French and Miller 1989). (4) Loss of (–) supercoiling. (5) Occlusion of *res* sites by other DNA binding proteins

Bacterial chromosomes and plasmids isolated from *E. coli* have roughly equivalent levels of $-SC$. The simplest comparative measure is supercoil density, σ , which is the ratio of the linking difference ΔLk of a supercoiled molecule to the linking number Lk_0 of a relaxed molecule under identical conditions. Supercoil density of plasmids can be established using agarose gel methods (Higgins and Vologodskii 2004), and the supercoil density of bacterial nucleoids can be measured by ethidium bromide titration in sucrose gradients (Pruss et al. 1986). Supercoiled DNA is an energized state. In eukaryotes almost all of the energy of supercoiling results from tight constrained looping of DNA around the histone octamer. In bacterial DNA, the value of σ has two components, constrained and unconstrained supercoil density $\sigma = \sigma_c + \sigma_u$ (Higgins and Vologodskii 2004). Proteins that constrain DNA curvature, or writhe, in vivo include the structural maintenance of chromosome (SMC) complex MukBEF, RNA polymerase, and the small nucleoid associated proteins (NAPs) (Schmid and Johnson 1991). Table 7.1 is a “back of the envelope” estimate of how much each characterized NAP contributes to constrained writhe. This list accounts for both the abundance of each protein and the defined topological characteristics that have been solved with in vitro assays or X-ray crystallography showing specific writhe of protein-DNA bound complexes. Thumbnail sketches of the small abundant non-enzymatic NAPs and the major catalytic enzymes that form nucleoids are given below. More detailed and quantitative information on each NAP is available in an excellent authoritative review of these proteins (Johnson et al. 2005).

Table 7.1 Nucleoid structural proteins^a

Protein	Mol./Cell in exponential phase	% Constrained – σ	Mol./Cell in stationary phase
RNAP	3,000	8%	3,000
HU	30,000	30%	15,000
IHF	6,000	6%	27,500
FIS	330,000	15%	1,000
H-NS	10,000	2–3%	7,500
StpA	1,000	Neg.	Neg.
Dps	5,500	?	10,000
MukBEF	100–500	?	?
Gyrase	500–1,000		500–1,000
Topo IV	500–1,000		500–1,000

^aConstrained supercoiling estimates for the most abundant nucleoid-associated proteins in *E. coli* in exponential phase. H-NS has a 10 bp DNA binding site and in vitro supercoiling experiments with Topo I yield one supercoil per 6 H-NS dimers (Tupper et al. 1994). StpA has not been measured in biochemical supercoiling assays, but its presence at 10% of the level of H-NS in exponential phase cultures and negligible amounts in stationary phase (C. J. Dorman, Personal communication) suggest a very minor contribution to constrained supercoiling. Neg. = Negligible

7.2 Catalytic Enzymes That Make or Move Supercoils in Bacterial Chromatin

A number of enzymes including helicases and DNA translocases utilize ATP to manipulate DNA topologically during normal processes of DNA metabolism. The enzymes mentioned below are keys to general nucleoid structure, either because of their abundance or their unique biochemical properties.

7.2.1 Topoisomerases

Four topoisomerases contribute to replication and segregation in most sequenced Gram-negative bacteria. Topoisomerases either break both strands of DNA simultaneously during the reaction cycle and are called type II enzymes or they only break one strand and are type I enzymes. In *E. coli* most evidence points to three topoisomerases playing dominant roles in controlling nucleoid function.

Topo I (the ω protein) is a type I enzyme and it is essential in *E. coli*. It was the first enzyme discovered that could change DNA linking number of plasmid DNA in vitro (Wang 1971). The characterization of the biochemical properties of Topo I led to the eventual understanding of the activity of all four topoisomerases that operate in Gram-negative bacteria (Cozzarelli and Wang 1990; Higgins and Vologodskii 2004; Postow et al 2001; Vologodskii et al. 2001; Wang 1991). Topo I either prevents R-loop formation by removing excess $-SC$, or perhaps facilitates the removal of R-loops from supercoiled DNA during transcription (Drolet et al. 2003).

DNA gyrase (Topo II) is the sole enzyme capable of introducing $-SC$ into covalently closed circular DNA. Gyrase is a hetero-tetramer consisting of two subunits each of GyrA and GyrB. Supercoiled DNA represents an energized state where the force required to form one additional supercoil is an exponential function of superhelix density. Gyrase can decatenate DNA, but its primary role is introducing $-SC$, which dissipates the $+SC$ driven by the unwinding of the Watson and Crick strands in front of replication forks. Gyrase is also required behind the fork to generate $-SC$ that condenses DNA for deposition into nucleoids. Gyrase is the primary enzyme for determining and maintaining a steady state value of σ in vivo. Two other topoisomerases contribute to σ (Khodursky et al. 2000; Zechiedrich et al. 2000). Relaxing activities of Topo I (ω protein) and Topo IV, the second type II enzyme, ensure that hyper-supercoiling does not occur (Menzel and Gellert 1983; Zechiedrich et al. 2000). Hypersupercoiling triggers formation of unusual DNA species including Z-DNA, R-loops, cruciforms, and H-DNA, which are all capable of interfering with normal cellular processes like transcription and DNA replication (Higgins and Vologodskii 2004).

Topo IV, the second type II topoisomerase discovered in Gram-negative bacteria, has structural homology with gyrase and is also a hetero-tetramer. It consists of ParC, which includes the DNA-binding domain (similar to GyrA) and ParE, which

has an ATP-binding domain related to GyrB. Topo IV was discovered in assays designed to find genes necessary for efficient F plasmid segregation. Although gyrase decatenates interlinked plasmids *in vitro*, Topo IV is the major decatenation enzyme in the cell (Espeli and Marians 2004).

Topo IV rapidly decatenates and unknots plasmids and it acts behind the fork to remove pre-catenane links in both plasmids and chromosomal DNA. Under some circumstances this enzyme can contribute to the topological balance of the chromosome (Espeli and Marians 2004). For example, a duplication of the Topo IV genetic region in *E. coli* can compensate for lethal mutations in Topo I that lead to hyper-supercoiling in otherwise WT strains (Raji et al. 1985). Topo IV is an essential gene in Gram-negative enteric bacteria, but this is not true for all bacteria. In *Mycobacterium tuberculosis*, DNA gyrase is the sole type II topoisomerase, so in this slow growing organism gyrase serves the dual function of supercoiling and decatenating the chromosome.

Topo III is the second type I enzyme found in bacteria. It functions as a decatenase to simplify DNA structure near the single stranded regions of replication forks (Hiasa and Marians 1994). Topo III is not thought to contribute to the *in vivo* –SC equilibrium.

7.2.2 RNA Polymerase

RNA polymerase ($\alpha_2\beta\beta'\omega$) with a molecular mass of 400,000 KD (Geszvain and Landick 2005) is the single enzyme that transcribes all classes of genes in most bacteria. However, this enzyme also accounts for a significant amount of constrained writhe in plasmids and chromosomal DNA. During exponential growth, about 2,000 molecules of RNA polymerase are engaged in transcribing genes around the genome (French and Miller 1989). Each transcribing enzyme unwinds a length of DNA equivalent to 1.7 supercoils in the form of a denatured segment of the DNA template (Gamper and Hearst 1982; Korzheva et al. 2000; Yin and Steitz 2002). The consequence is that RNA polymerase contributes roughly 3,400 supercoils, or 10% of constrained writhe. The distribution of these enzymes along the genome is non-random, so this component of supercoiling is highly restricted (see below and Fig. 7.4).

7.2.3 The Replisome

DNA replication creates the largest topological challenge to untangling bacterial DNA *in vivo*. Replication involves assembly of two large protein machines called replisomes that contain about 30 individual proteins (O'Donnell 2006). Each replisome synthesizes DNA at the astounding rate of 500 nucleotides/s. As replisomes pass, all protein must dissociate and the interwound Watson and Crick strands are

separated by the DnaB helicase. There are two supercoil problems to solve. First, in front of the fork, the helicase creates +SC that must be dissipated to allow fork movement. Second, behind the fork, –SC must be re-introduced to promote condensation and establish the proper ratio of constrained and unconstrained σ in two new nucleoids.

7.2.4 *MukBEF*

MukBEF is a large complex composed of three proteins, MukB, MukF, and MukE. MukB is a large 171 kDa protein in the structural maintenance of chromosome (SMC) family of proteins (Cortes-Ledesma et al. 2007). Like all members of the SMC family, it contains an N-terminal globular domain with a Walker A ATP-binding motif, an alpha helical region, a hinge domain, another alpha helical region and a C-terminal globular domain with a Walker B ATP-binding motif. The protein forms anti-parallel dimers with the two N- and C-terminal globular domains forming a head. Therefore, the Walker A and B ATP binding site is formed by complementary subunits. Each head contains a DNA-binding site and an interaction site for the MukF protein. The Muk EF complex contains a dimer of MukF and each MukF subunit binds two MukE proteins (MukF-(MukE)₂)₂. The simplest MukBEF complex contains one MukB dimer with each headpiece interacting with one MukEF complex. ATP binding and hydrolysis promotes the formation of larger assemblies in which each MukB head interacts through MukEF ligands with other MukB subunits (Woo et al. 2009). In eukaryotes, SMC proteins carry out essential roles as chromosomes pass through the mitotic cycle. The eukaryotic family of SMC proteins includes condensins I and II, which compact chromosomes, and cohesin, which binds replicated sister chromosomes together until they align on the metaphase plate and are pulled to opposite poles in anaphase. MukBEF is not essential in many bacteria, but it plays a role that is similar to eukaryotic condensins (Sawitzke and Austin 2000; Espeli and Mariani 2004; Weitao et al. 2000). MukBEF compacts bacterial DNA and induces specific right-handed chiral knots when binding is done in the presence of a type II topoisomerase (Petrushenko et al. 2006). High-resolution structures suggest that DNA gains access to the interior portion of these molecules with conformation changes in the MukEF proteins during the ATP hydrolytic cycle (Woo et al. 2009).

7.2.5 *FtsK*

FtsK is an essential protein encoded by the *ftsK* gene. The N-terminal part of FtsK is an integral membrane protein that is anchored in the Fts complex at the septum. A large C-terminal domain (FtsK_C) forms a hexameric ring that can bind to dsDNA and hydrolyze ATP. An FtsK_C deletion is viable, but the C-terminal domain of the

enzyme can be synthesized independently and it has been characterized biochemically. FtsK_C contains an AAA + ATP binding site and is an efficient dsDNA translocase that moves along DNA at the rate of 7 KB/s (Bigot et al. 2007). The enzyme is required for XerCD catalyzed recombination, which converts circular dimer chromosomes, which are generated by homologous recombination, to monomers by site-specific recombination at the *dif* site (Lesterlin et al. 2004). The FtsK_C protein also stimulates the decatenation activity of Topo IV at the *dif* site (Espeli et al. 2003), and it contributes to removal of catenane links in cells deficient in Topo IV using the topoisomerase activity of the XerCD protein at *dif* (Grainge et al. 2007).

7.3 NAPs of Bacterial Chromatin

Because about half of bacterial chromosome supercoiling is associated with proteins, the term histone-like has been used to describe some of these elements. However, they clearly have no relationship with true histones and their binding mechanisms are varied and much different from each other as well as from histones.

7.3.1 HU

HU protein is a heterodimer of the HupA and HupB proteins that is present at about 15,000 copies/cell (Ali Azam et al. 1999). In addition to the heterodimers, HupA and HupB form homodimers, but the significance of homodimers in WT cells is not known. HU binds a 36 bp length of DNA, and it strongly bends DNA (Johnson et al. 2005). It also binds to ssDNA and RNA. HU constrains supercoils in vitro, and produces highly –SC DNA when protein is incubated with a relaxed plasmid in the presence of an enzyme like calf thymus Topo I, which removes either positive or negative supercoils. Based on the observed change in supercoil density in strains defective in HU in both *E. coli* and *Salmonella* (Hillyard et al. 1990; Huisman et al. 1989), HU is the NAP that makes the largest single contribution (30–40%) to constraining –SC.

7.3.2 IHF

IHF is structurally related to HU protein. It is composed of a dimer of IhfA and IhfB proteins and, like HU, has a 36 bp dsDNA-binding site (Johnson et al. 2005). IHF bends DNA by about 150° when a 13 bp consensus sequence of WATCAANNNTTTR is present. DNA bending by IHF is critical in the phage λ integration reaction, and IHF binding near promoters strongly influences expression of about 100 genes in *E. coli* (Long et al. 2001).

7.3.3 FIS

FIS is a homodimer of the *fis* gene product in both *E. coli* and *Salmonella*. FIS is most abundant when ribosomal RNA and protein production requirements are the greatest, i.e., when cells are growing exponentially in rich medium at 37°. It is dramatically down regulated in stationary cells. FIS can supercoil DNA weakly in the Topo I plasmid supercoiling assay and enhances reactions like the phase inversion switch in the flagellin genes in *Salmonella* (Johnson and Simon 1985). Measured DNA bending angles for FIS vary from 45° to 90° using different specific sites. FIS has also been observed to form DNA loops in single molecule experiments when the protein is applied to DNA in high concentrations (Skoko et al. 2006).

7.3.4 H-NS

H-NS is a homodimeric protein that forms higher order complexes in solution. H-NS binds dsDNA, ssDNA and RNA. It binds dsDNA with a 10 bp footprint, but it can cooperatively spread along DNA and has been implicated in regulating over 300 genes in *E. coli*, *Salmonella*, and other organisms (Castang et al. 2008). When expressed at high levels, H-NS can inhibit nearly all RNA transcription and induce an artificial stationary phase (McGovern et al. 1994). Whether H-NS constrains supercoils in vivo is not clear. In vitro experiments indicate that it supercoils plasmid DNA in the presence of Topo I at a level of 1 supercoil per 12 monomers (Tupper et al. 1994). In vivo experiments indicate modest effects of the null H-NS mutations on chromosomal supercoiling (Hardy and Cozzarelli 2005). In plasmid experiments, reports vary from H-NS null mutations causing slight increases in –SC, slight decreases in –SC, to no change. Sequence probably plays a large role in what types of structure H-NS forms (Doyle et al. 2007). It has been observed to stabilize an interwound form of supercoiling (Dame et al. 2000) rather than the solenoidal supercoiling associated with HU. Under high protein concentrations, it can also form a bridge between two DNA molecules in vitro (Dame et al. 2006).

7.3.5 StpA

StpA is a protein with significant homology to H-NS. It also binds DNA with a 10 bp footprint and forms homodimers as well as heterodimers and higher order complexes with H-NS. Curiously, this gene was discovered as an *E. coli* factor involved in RNA splicing of a bacteriophage protein (Zhang and Belfort 1992), so it may have a stronger role in RNA metabolism than in DNA condensation. Estimates vary with strains and technique, but a recent estimate is that StpA is present at only 10% of the level of H-NS (C. J. Dorman, Personal Communication.) Most H-NS phenotypes appear unaffected by mutations in StpA.

7.3.6 *Dps*

Dps is a protein with relatively little impact on chromosome structure in actively growing cells. However, as cells enter stationary phase, this protein becomes the predominant nucleoid bound factor and it forms tightly packed arrays (Minsky and Kolter 2005). One potential role for Dps is to bind iron, which protects the chromosome from oxidative damage. NAPs and their architectural properties are discussed in larger detail in Chapter 8.

7.4 Nucleoid Domains

Numerous experimental systems have shown that there are regions of the chromosome with special properties that are often called domains. This term domain may describe specific regions with exceptional exposure to or protection from physical or enzymatic processes. For example, the region of about 100 kb centered around the *dif* site is known as the DAZ (*dif* activity zone) for its exceptional properties and exposure to the FtsK DNA translocase (Louarn et al. 2005). Some transposons generate a field of protection called the immunity zone that shields the element from inserting into itself or into other similar transposons (Manna and Higgins 1999). Studies using the lambda Int protein have shown that five or more segments of the *E. coli* and *Salmonella* genomes may have macro-domain status that make inter-domain interactions more efficient than cross domain interactions (Garcia-Russell et al. 2004; Valens et al. 2004). And both *E. coli* and *Salmonella* contain regions in the chromosome and in the large plasmids including F and pSLT respectively that are insulated from or hyperactive for transposition by bacteriophage Mu (Manna et al. 2004; Manna et al. 2007). But the mechanism(s) that creates and maintains supercoiling domains has been one of the most challenging mysteries to solve.

7.5 Supercoil Diffusion

The unconstrained component of $-SC$ in bacteria represents about half of the total value of σ in bacterial cells (Bliska et al. 1991; Jaworski et al. 1991; Pettijohn and Pfenninger 1980). In 1980 Sinden and Pettijohn reported a scheme to measure unconstrained supercoils by monitoring torsional effects of X-ray-induced DNA chain breaks in vivo (Sinden et al. 1980). They combined X-ray treatment and UV-induced psoralen crosslinking to study the supercoil dynamics of chromosomes. This technique exploits the fact that psoralen intercalates more readily into torsionally strained $-SC$ DNA than it does into relaxed DNA. By quantitatively measuring psoralen-DNA adducts formed as a function of X-ray dose, they estimated the number of breaks needed to relax an entire chromosome. Their calculations indicated that the *E. coli* chromosome had 40 domains (Pettijohn 1996;

Sinden and Pettijohn 1981). However, their model assumed that all domains were of equal size and that the probability of X-ray hits in every domain would be equivalent. These assumptions turned out not to be correct.

In 1996 a new assay permitted the analysis of $-SC$ diffusion inside living cells unperturbed by physical or chemical modification (Higgins et al. 1996). By introducing 114 bp *res* sites into specific locations in the *Salmonella* chromosome and then inducing pulsed expression of $\gamma\delta$ resolvase, one can measure the diffusion of plectonemic supercoils within discrete segments of the chromosome. If two *res* sites form a synapse by trapping three negative supercoiled nodes, then a recombination reaction deletes the intervening DNA sequence (Higgins et al. 1996). This reaction is depicted in Fig. 7.1. On the left, the 114 bp *res* site includes three closely spaced sub-sites (*res*-I, *res*-II, and *res*-III); each of these binds a dimer of Resolvase. Once Resolvase is bound, movement of negative supercoiled DNA is required to form an interdigitated tangle of each sub-site in a synapse that traps three negative node crossings of duplex DNA. Only after this structure forms will the enzyme break and rejoin four phosphodiester bonds within *res*-I, producing a deletion having the recombinant DNAs singly interlocked as a supercoiled catenane.

DNA conditions that block the resolvase synapse formation are listed on the right side of Fig. 7.1. First, if there are loops in the DNA (Fig. 7.1(1)) that involve binding proteins at a loop base, recombination will be restricted to *res* sites within each loop (Saldanha et al. 1987). The reason is that the slithering and branching of the supercoiled network cannot form a three-node tangle when two *res* sites are on different DNA molecules or are in different loops of the same DNA molecule. Second, tangling of sister DNA strands about each other can impede or block slithering and branching (Fig. 7.1(2)). Intertwining of DNA strands behind a replication fork generates a specific tangle pattern that is often referred to as pre-catenanes. These entanglements link partially replicated sister chromosomes and block segregation (Adams et al. 1992). Third, any segment of chromosomal DNA with a high-density of bound proteins may block resolution. The case in Fig. 7.1(3) shows the high density of RNA polymerase molecules transcribing a ribosomal RNA operon (French and Miller 1989). Fourth, any condition that causes loss of supercoiling blocks synapsis. This includes DNA breaks and enzymatic relaxation of negative supercoils Fig. 7.1(4). Fifth, any protein that occludes the resolvase from binding to one of the *res* sub-sites will prevent recombination Fig. 7.1(5).

7.6 Supercoil Diffusion in the Nucleoid

The favored model of chromosome structure predicted by the results of Sinden and Pettijohn was an array of equal sized loops of chromosomal DNA (Fig. 7.2 left). If one introduces *res* sites systematically along the genome so that recombination tests the intervening region for barriers, the model predicts a site-specific pattern. As one tests the chromosome over longer and longer distances, there should be sites

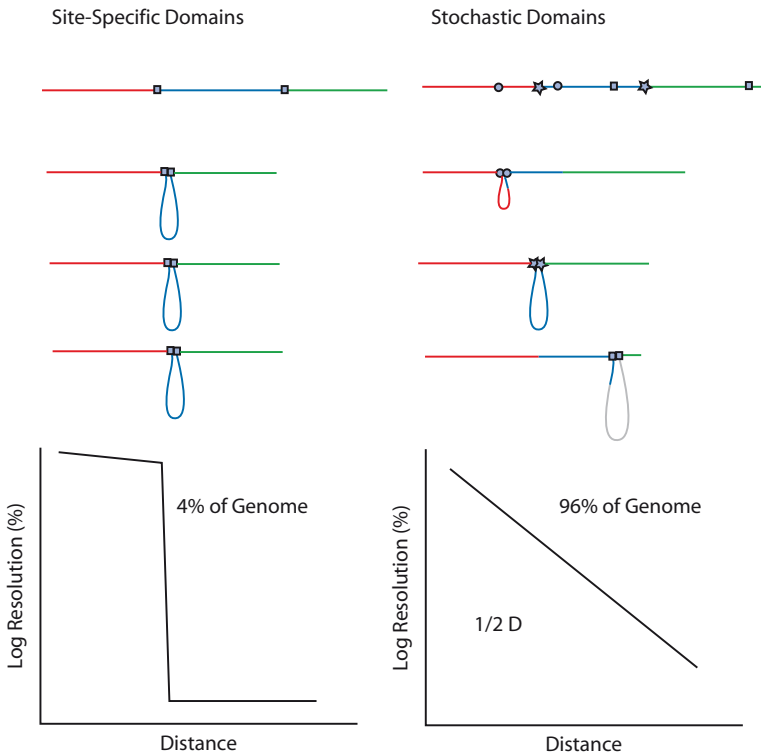


Fig. 7.2 Two models of domain boundary structure. On the left, site-specific barriers generate a genome with a fixed domain length from cell to cell. Resolution reactions should decrease with dramatic barrier effects. The stochastic model on the right posits that stochastic domains generate domains that vary in both position and in domain length for a given segment of the genome. Site-specific recombination decays as a first order process with a term called the $1/2D$ being the distance over which half the cells have a barrier. The stochastic model is true for most of the genome in *E. coli* and *Salmonella* (see text for explanations)

that dramatically inhibit recombination of *res* sites located on opposite sides of a barrier. The alternative model is stochastic barrier location (Fig. 7.2 right). In this case, barriers are located by a statistical rule, perhaps a Poisson distribution; systematic assays over increasing distance would exhibit a first order decay (Higgins et al. 1996). Resolution assays spanning 2% of the *Salmonella* genome supported stochastic organization rather than the sequence-specific model (Fig. 7.2) (Staczek and Higgins 1998; Stein et al. 2005).

The Cozzarelli lab provided a confirmation of the stochastic model using an independent genome-wide approach in *E. coli*. First, microarray experiments were culled to compile a list of 300 genes distributed throughout the genome that consistently and reliably respond to supercoil changes. About 100 genes gave reliable expression **increases** when the chromosome was relaxed, and 200 genes reliably **decreased** expression within 5 min of DNA relaxation (Peter et al. 2004).

Postow and colleagues exploited a system for introducing breaks at defined positions by expressing restriction enzymes under tightly controlled conditions. They measured the global gene expression response to chromosome relaxation after introducing site-specific breaks and combined it with information on the distance of each promoter from a restriction site. Four models were tested using Monte Carlo simulations to find the best fit to the genome transcription data. The models were: (a) fixed barrier sites and fixed domain length (Fig. 7.2 left), (b) variable barrier sites and fixed domain length, (c) fixed barrier location and variable domain length, and (d) variable barrier location and variable domain length (Fig. 7.2 right). Their conclusion was that the *E. coli* genome consists of variable sites and variable domain lengths throughout most of the genome. Consequently, two experimental techniques carried out in two well-characterized clades of bacteria agree that chromosomes have a predominantly stochastic domain structure; the *E. coli* and *Salmonella* genomes are partitioned into approximately 500 domains with a median size of 10 kb (Deng et al. 2005). Thus, rather than having an organization based on specific domains all equal in size, bacterial chromosomes are like snowflakes – each one is different from another.

7.7 Supercoil Dynamics of Transcription

In addition to stochastic domains covering most of chromosomal DNA, a subset of regions contains site-specific domain boundaries. These regions are highly transcribed genes that occupy about 4% of the physical sequence. In 1999, Scheirer first reported that transcription induced by de-repression of the Mu early promoter in a lysogenic strain of *Salmonella* caused a supercoil diffusion barrier to appear within proviral DNA (Scheirer and Higgins 2001). To follow up, Deng generated test intervals at several locations in the *Salmonella* chromosome where transcription could be efficiently turned on and off (Deng et al. 2004). She discovered that constitutive expression of TetA protein, which is the Tn10 membrane-associated pump that keeps tetracycline concentrations low in vivo, caused the appearance of a transcription domain in an otherwise normal region of the chromosome. Subsequent tests using synthetic modules controlled by the Tn10 TetR repressor protein showed that high level transcription of essentially any gene or operon, including ones encoding cytoplasmic proteins like β -galactosidase (β -Gal) and aminoglycoside-3'-O-phosphotransferase, reduced resolution by more than tenfold. In every case, barrier(s) to resolution disappeared when transcription was turned off.

7.7.1 What Does a Transcription Domain Look Like?

X-ray crystal structures show that DNA enters RNA polymerase through one channel and RNA emerges from another (Geszvain and Landick 2005; Perederina et al. 2004). Thus, transcription must strip DNA of all attached proteins, or push them

along the DNA out of the path of RNA polymerase. After the first enzyme passes, an impediment free zone allows efficient RNA movement throughout the transcribed region. Transcription barriers were observed for both artificial constructs as well as for highly transcribed genes tested in their normal chromosome context (Deng et al. 2004).

To further characterize the topology associated with high transcription, tests were made using resolution modules placed at multiple locations around a synthetic construct in the chromosome. Four rules were apparent. First, high transcription of any single gene or multi-gene operon that lies between a pair of *res* inhibited resolution of the highly transcribed sector. Loop models for domain behavior are popular (Cook 2003) and a loop model, as resulting from transcription, is illustrated in Fig. 7.3.

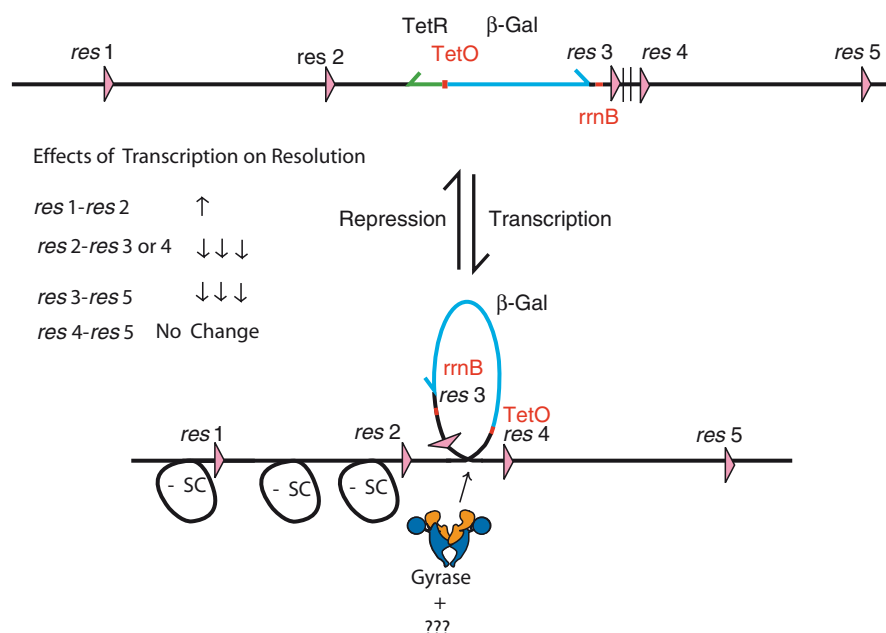


Fig. 7.3 Transcription-induced domain structure of a synthetic operon in which the β -Gal protein is regulated by the TetR repressor and transcription stops at the strong *rrnB* transcription terminator. Addition of inducer causes strong inhibition (down arrows) of resolution involving all sites flanking the transcription unit. Resolution of sites *res 1* and *res 2* that both lie upstream becomes elevated (up arrow) by transcription. Recombination involving *res 3*, which is close to the terminator, is impaired with both upstream (*res 2*) and downstream (*res 5*) sites. At *res 4*, which is 500 bp beyond the terminator, there is no effect of transcription on resolution reactions involving *res 5*

7.7.2 *Why Does Transcription Block Site-Specific Recombination?*

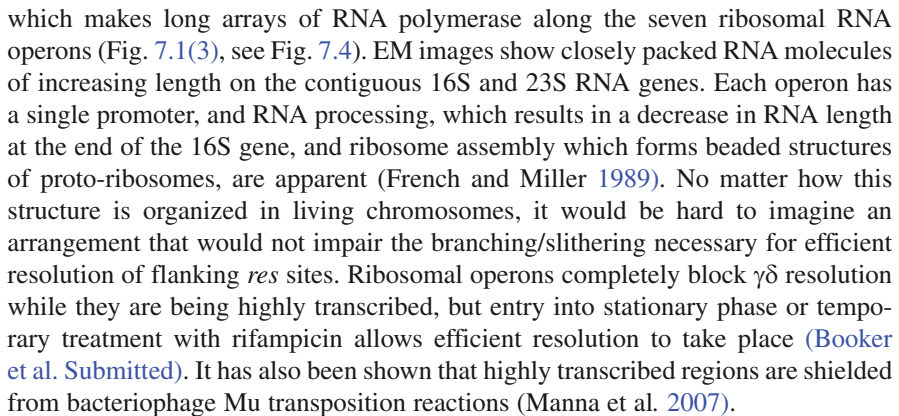
First, Deng demonstrated that turning transcription on and off made the domain appear and disappear, so the model has two states, looped and unlooped respectively. Second, recombination between *res* 1 and *res* 2, which both lie upstream of the transcription, actually increased during transcription. This is consistent with Liu and Wang's proposed twin domains of supercoiling hypothesis for transcription (Liu and Wang 1987). The twin domain model assumes that there is a significant rotational frictional coefficient associated with highly transcribed genes, and that the movement of the static polymerases and associated transcripts generates negative supercoils upstream from the transcription unit coupled with positive supercoiling downstream of the transcribed region. Increased negative supercoiling would predict stimulation of $\gamma\delta$ recombination in the upstream region, and that prediction is observed. Although information on supercoil dynamics is scant in specified regions of bacterial chromosome, data from plasmid experiments using structures like Z-DNA that form under high levels of supercoiling are consistent with elevated $-SC$ in the region immediately upstream from promoters (Rahmouni and Wells 1992).

Third, at *res* 3, located 150 bp downstream of the transcription terminator, transcription inhibits recombination with both the *res* 2 site upstream and the *res* 5 site downstream see Fig. 7.2 (Booker et al. Submitted). One interpretation of this result is that the transcription loop includes the terminator (*rrnB*) plus a little extra DNA (Fig. 7.2). An alternative explanation is that some protein binds the region downstream of the *rrnB* terminator and occludes binding of Resolvase to sub-sites in *res* 3. One candidate for occluding the *res* 3 site is DNA gyrase, which preferentially binds to relaxed DNA (Higgins and Cozzarelli 1982). By supercoiling DNA downstream of highly transcribed regions, gyrase could eliminate $+SC$, or perhaps more likely, ameliorate decreases in $-SC$.

Fourth, since resolution requires negative supercoiling, the fact that there is no inhibition of resolution during transcription between sites *res* 4 and *res* 5 places a 500 bp limit on the length of DNA that might lose $-SC$ when the promoter is active. The gyrase footprint on DNA is about 150 bp of DNA at the high affinity site from bacteriophage Mu (Oram et al. 2006), so gyrase could also occlude binding to *res* 3. Significantly, global analysis of gyrase in the *E. coli* genome shows that the distribution of gyrase co-localizes strongly with the highly transcribed genes (Jeong et al. 2004).

7.7.3 *Do the Data Prove Looping?*

Loop models are not the only explanation for these data. Regions of DNA that contain high densities of DNA binding proteins might restrict resolution by other mechanisms. One example of a high-density protein zone is RNA polymerase traffic,



7.8 Conservation of Nucleoid Transcription Patterns Between Species

The locations of domains that contain reliable sequence-specific barriers to $\gamma\delta$ resolution due to high-level transcription are highly conserved in the DNA sequences of *E. coli* and *Salmonella typhimurium* (Fig. 7.4). Microarray data for cells grown in rich LB medium show that over 75% of all genes have a steady state RNA abundance relative to DNA of less than 1. Thus, cells with no RNA molecules for these genes are not rare. Genes in the low RNA abundance class include 30% of the essential genes, so that high RNA abundance is not a prerequisite for being essential. On the map, the genes that are transcribed at the highest rate are the seven ribosomal RNA operons, which are shown in blue. Seventy percent of all the RNA polymerase molecules are engaged in ribosomal RNA transcription under maximal growth conditions (Bremer and Dennis 1996). Each gene listed in black encodes a protein and a number that represents the steady state RNA/DNA ratio for the gene transcript. Significantly, although 35% of the orfs in *E. coli* or *Salmonella* are not present in the other species, 100% of the top 100 transcribed genes are present in both organisms, and the transcription levels are quite similar. Thus the highly transcribed genes are under strong selection for both function and expression level. The striking maintenance of gene organization among the highly transcribed genes is only broken by an inversion around the terminus region (*dif*). However, even in the inversion, the gene order has been maintained within the inverted segment.

7.9 Post-DNA Replication Nucleoid Formation; Slow and Fast Doubling Times

Nucleoid assembly and association of nascent DNA with the various NAPs must occur quickly after replisomes generate sister strands. Most well studied bacteria have a single large chromosome with a DNA initiation region called *oriC*. At *oriC*, two replisomes are assembled and they proceed bi-directionally following a path from *oriC* to a point approximately 180° on the opposite side of the chromosome. Termination of replication occurs near the *dif* site in *E. coli* and *Salmonella* (Kuempel et al. 1991). One replisome duplicates replicore I and the other copies replicore II. Each replisome synthesizes DNA in a semi-discontinuous mode. One new strand grows into a 5'–3' polymer in the same direction as fork movement, and synthesis of this DNA is continuous, or rarely interrupted by discontinuities from start to finish. The complementary strand is synthesized discontinuously with multiple initiation events that generate short Okazaki fragments, which are joined together by ligase after removal of RNA primers on each molecule.

7.9.1 Nucleoid Formation in Cells Growing on Defined Medium

At slow growth conditions, synchronous cultures of *E. coli* cells made using a “baby machine” method (Bates et al. 2005) with a doubling time of 125 min exhibit eight important time signatures (Fig. 7.5). (1) DNA replication begins at 17 min after elution from the column and DNA synthesis continues for 50–60 min. (2) For 10–14 min period after replication begins, the duplicate copies of *oriC* remain together (Bates and Kleckner 2005). (3) At 27–32 min, the two *oriC* regions move to the $\frac{1}{4}$ and $\frac{3}{4}$ positions of the cell, which is the position that will become cell centers in the new daughter cells. (4) Nucleoids can be visualized in a special medium where the refractive index is carefully controlled, or by staining with DAPI. In baby machine conditions, nucleoids split at about 60 min. (5) Replication is complete by 67–77 min. (6) Separation of fluorescent markers near the terminus occurs at about 80 min. (7) At 100 min closure of the septum begins. (8) Finally

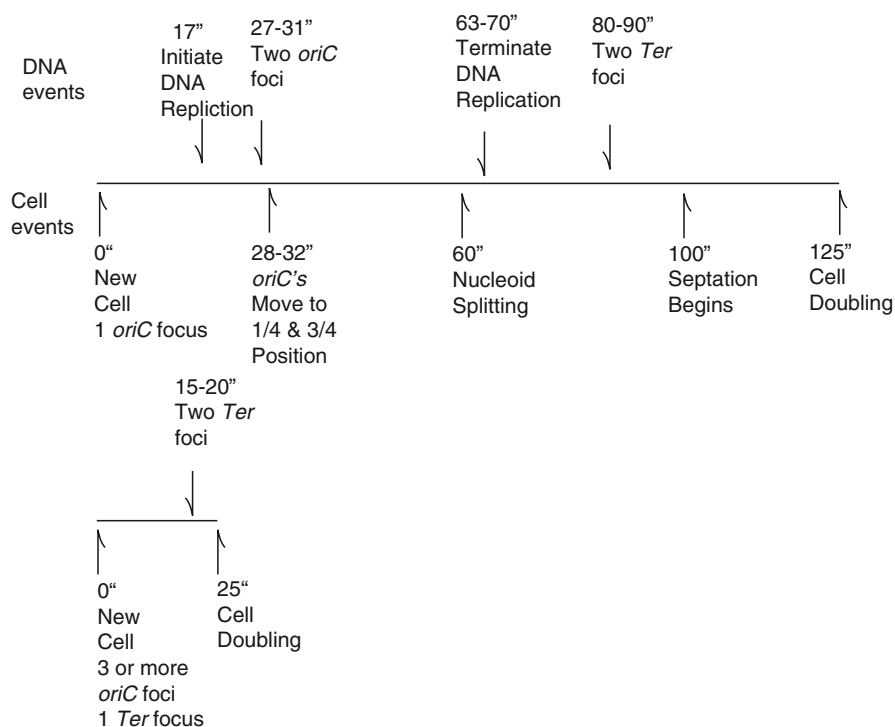


Fig. 7.5 Time signature of discernable events in an “average” cell growing exponentially in minimal medium with a 125 min doubling time (*top*) or in rich LB medium with a 20 min doubling time (*bottom*). Data abstracted from (Bates and Kleckner 2005; Nielsen et al. 2006; Wang et al. 2006; Wang et al. 2008)

cell division is completed 40 min later at 120 min. A diagram of nucleoid division patterns is shown in Fig. 7.6. In newborn cells there is a single MukBEF focus for slow growing cells. During the synthetic period, the focus doubles and each moves

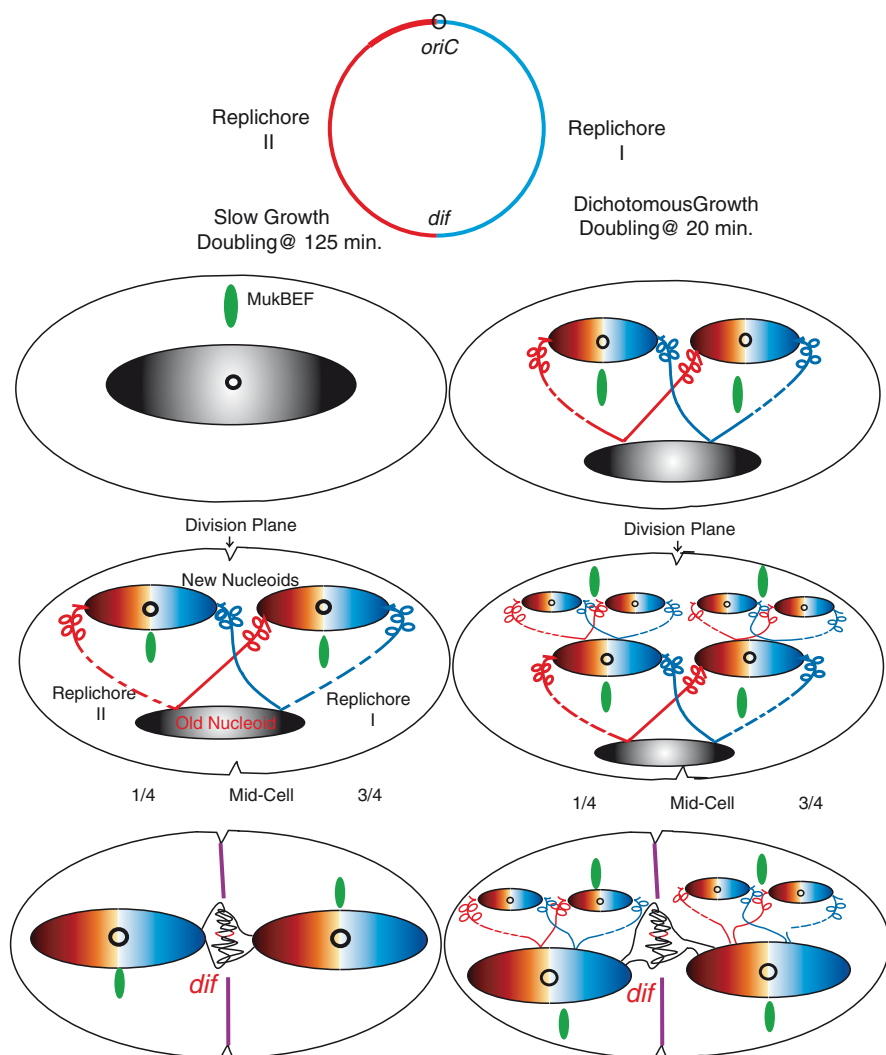


Fig. 7.6 Chromosome organization and segregation patterns during slow growth (left) and dichotomous growth (right). The DNA replication pattern from *oriC* to the terminus near *dif* is shown at top with replicore I blue and replicore II red. Replication of the continuously synthesized and discontinuously synthesized strands from each replicore are shown as solid and broken lines, respectively. MukBEF condensin foci are indicated as green ovals near the quarter positions of the long cell axis. The cell septum is marked by a purple line at mid-cell

to the quarter positions of the cell, i.e. between midcell and cell poles on the long cell axis.

The structure and binding force(s) that stabilize *oriC* near the quarter positions are not understood, although the MukBEF condensin plays a part. Estimates on the copy number of MukBEF molecules differ from 100–500 per cell. MukBEF forms a focus at the quarter positions (shown as green ovals in Fig. 7.5) (Niki et al. 2000), and this complex helps condense nucleoids into structures that segregate efficiently (Sawitzke and Austin 2000). A *mukB* null mutant fails to position *oriC* at the quarter cell locations, which results in the formation of nucleoids with a more random organization than WT nucleoids (Danilova et al. 2007).

Nucleoid development has been followed in a number of different labs using *E. coli* strains tagged with Lac and Tet operator arrays decorated with two colored fluorescent repressors or *parC* sites decorated with fluorescent ParB protein. The following summary of events represents a hypothetical average cell in an asynchronous population (Fig. 7.6). The continuous strand from one replichore (here arbitrarily labeled replichore I) is assembled into chromatin and deposited at the inner margin of the left nucleoid. The discontinuously synthesized strand of replichore I, shown as broken blue line, is deposited at the outer margin of the right nucleoid. The continuous and discontinuous strands from replichore II move to the inner and outer edges of the right and left nucleoids respectively. *E. coli* generates nucleoids that contain genes from different replichores lying bilaterally on different sides of the origin. Genes closest to the origin are followed by genes positioned toward the terminus. This leads to a direct replichore repeat pattern in most cells (63%) with a minor fraction of cells (30%) becoming mirror image replichore nucleoids (Wang et al. 2006; Wang et al. 2005). This pattern is distinctly different from *Caulobacter crescentus*, which generates exclusively mirror image nucleoids with *oriC* at one pole and *dif* at the pole that has either a stalk or a flagellum (Viollier et al. 2004).

Both early and late stages of segregation appear to take extra time in slow growth conditions. The delay between initiation and movement of the *oriC* to relatively fixed positions has been called cohesion, and the mechanism may be complex (Bates and Kleckner 2005; Sunako et al. 2001). One possibility is that pre-catenane links between the sisters (Fig. 1.2) requires time to organize and untangle this region to allow freedom to move to quarter positions (Grainge et al. 2007). There is also a significant delay between the completion of DNA replication and separation of fluorescent foci near the terminus, which may be linked to the MatP system. MatP binds specific sites and organizes a domain near the terminus (Mercier et al. 2008). Complete decatenation of links between sister chromatids, is necessary for complete segregation, and this process is catalyzed most efficiently by Topo IV, which uses the DNA binding and motor domain of FtsK to do the job efficiently.

About 15% of the cells undergo RecA-dependent homologous recombination that generates a dimer chromosome. Chromosome dimers are resolved by the combined action of XerC, XerD, and FtsK working at the *dif* site. Complete decatenation is carried out by the combined action of gyrase, Topo IV, FtsK, XerC, XerD and polarized G-rich Kops sequences that are polarized about the *dif* site.

7.9.2 Nucleoid Formation During Dichotomous Replication

Some bacteria have evolved an adaptive ability to compress cell division stages to achieve a very rapid doubling time. Rapid growth is a strong selective advantage during the competition for a niche in the mammalian gut (Freter et al. 1983). Dichotomous replication drastically constricts the time period for carrying out all the functions that are carried out in sequential order during slow growth (Fig. 7.5, lower). Dichotomous growth depends on having a large biosynthetic capacity so that cells can synthesize proteins, nucleic acids, and all cellular components at break-neck speeds needed to produce viable cells every 20 min. Organisms that carry out dichotomous growth exhibit a signature of multiple copies of ribosomal RNA operons. Sueoka described dichotomous growth first in *B. subtilis* (Sueoka 1971), which has 12 *rrn* (ribosomal RNA) operons. Both *E. coli* and *Salmonella* grow dichotomously and have 7 *rrn* operons (Fig. 7.4). By contrast, *Caulobacter crescentus*, which doubles slowly in over an hour, has only 2 *rrn* operons to satisfy its protein requirements, and the very slow growing *M. tuberculosis* with a 12 h doubling time requires only one *rrn* operon.

Chromosome segregation becomes more complicated under dichotomous growth conditions. First, each newborn cell has two MukBEF foci, and during the division cycle 3–4 foci are seen. When a critical mass to volume is reached in *E. coli*, and replisomes have moved a significant distance toward the terminus, cells synchronously reinitiate new forks using both replicated *oriC* sequences. This adds four additional nucleoids to the picture (Fig. 7.5 right side.) The segregation pattern is now complicated and the timing of each step is constricted and difficult to sort out using marked chromosomes in live cells. Nonetheless, at the time of cell division, each daughter cells must untangle one complete sister chromosome and 2 (or even 4) partially replicated arcs with bi-directional replicating *oriC* elements.

7.10 Dichotomous Chaos at *dif*

All barriers to the replicative DnaB helicase must be eliminated to allow unimpeded fork movement at the normal rate of >500 bp/s. DNA replication in cells carrying the *gyrB652* mutation bogs down due to the sluggish supercoiling action of a hypomorphic gyrase (Pang et al. 2005). This impediment mimics the effect of gyrase inhibited by quinolones, which impedes RNA polymerase (Willmott et al. 1994) and causes fork arrest when DNA polymerase encounters the gyrase-quinolone-DNA complex (Hiasa and Marians 1996). Once a fork becomes arrested, there is a critical time for rescue before it is overtaken by the next round of synthesis. A second fork running into a stalled fork causes chromosome breaks (Bidnenko et al. 2002). The experimental signature of dichotomous chaos in *Salmonella* is a dramatic (30-fold or more) loss of $\gamma\delta$ -catalyzed resolution near the *dif* site (Pang et al. 2005). This loss in recombination potential is likely caused by unresolved catenane links that make synapsis inefficient, and possibly by a loss of –SC around the *dif* site caused by replication fork convergence.

In addition to death caused by a sluggish gyrase, a similar phenotype is triggered in null mutants of MukB (Niki et al. 1991) or by hyper-initiation. Grigorian et al. showed that over-production of DnaA leads to over-initiation at *oriC* and fork collapse (Grigorian et al. 2003). More extensive analysis by Simmons et al. demonstrated that hyper-initiation of DNA replication in *E. coli* leads to breakage of chromosomal DNA (Simmons et al. 2004). They analyzed fork movement with microarrays but found no specific sites of fork arrest. They suggested that merging forks could cause breaks under hyper-initiation conditions. This hypothesis also explains the sickness of *seqA* mutations. These strains are only healthy in minimal medium at low temperature because they lack the sequestration mechanism in which SeqA binding inhibits premature reinitiation by blocking the DnaA protein's access to *oriC* (Campbell and Kleckner 1990; Skarstad et al. 2001).

7.11 Differences in Species-Specific Supercoil Set Points

Supercoiling is a critical factor for nucleoid formation and segregation. The multiple functions of negative supercoiling include: (1) Gyrase-dependent compaction of DNA into an interwound and branched plectoneme. (2) Segregation of sister chromosomes requires unknotting and decatenation of newly synthesized double strands. All simplifications in DNA topology are dramatically stimulated by negative supercoiling, which brings the interwound strands of DNA close together and makes inter-chromosomal links more apparent to the decatenase, Topo IV (Rybenkov et al. 1997). Single molecule experiments show that Topo IV acts preferentially on positive nodes rather than on the negative nodes of $-SC$ DNA (Stone et al. 2003). (3) Over 300 genes, including the promoters for *gyrA*, *gyrB*, and *topA*, are regulated by supercoiling stress (Peter et al. 2004). Considering the central importance of gyrase in nucleoid formation, it was surprising to discover that WT strains of *E. coli* and *S. Typhimurium* generate substantially different levels of in vivo torsional strain. Nonetheless, plasmid pBR322 extracted from *E. coli* has a median σ value of -0.70 whereas the same plasmid from *S. Typhimurium* has a value of -0.06 (Fig. 7.7) (Champion and Higgins 2007). As one illustration of the significance of this difference, WT *E. coli* is not viable when grown at the *Salmonella* level of σ .

7.12 The Paradox of Supercoil Dynamics in *E. coli* and *Salmonella*

Why have *E. coli* and *Salmonella* evolved to maintain different levels of σ ? How do two organisms regulate >300 supercoil-responsive genes at different levels of torsional strain? How does nucleoid formation differ between two species with such large differences in average supercoiling? Answers to both questions are not

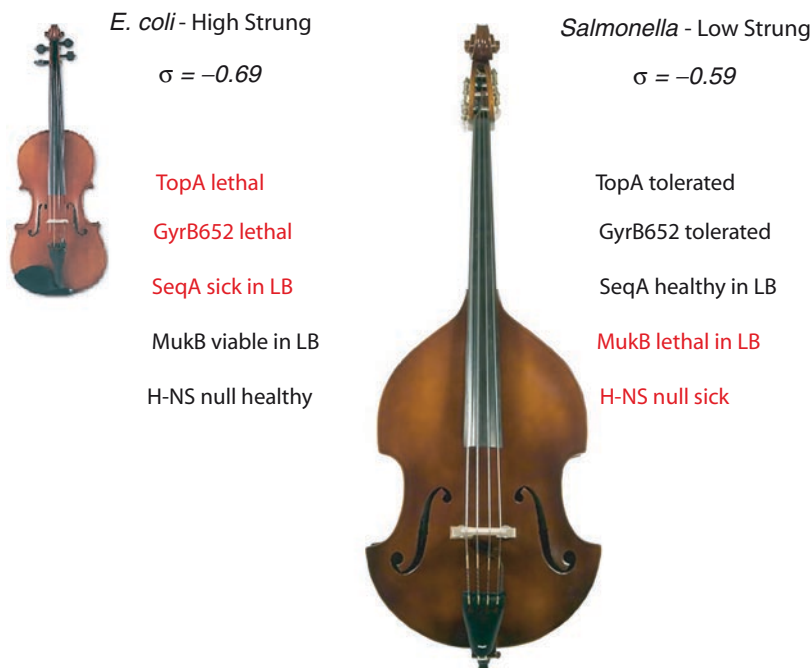


Fig. 7.7 Comparison of supercoil density and nucleoid related phenotypes of *E. coli* and *Salmonella typhimurium*. Red = severe to lethal phenotype, Black = mild to non-detectable phenotype

entirely solved, but it is clear that these organisms have many significant phenotypic differences for identical mutations in topologically sensitive genes. The list includes *gyrB*, *topA*, *seqA*, *mukB*, and *hns* see Fig. 7.7 (Champion and Higgins 2007; Falconi et al. 1991; Navarre et al. 2006). Although there could be complicated explanations that cannot be excluded in each case, supercoiling provides a simple unifying explanation for all of the observed phenotypes.

Amber mutations in *topA* (ω protein) have long been known to be lethal in *E. coli* but are healthy in *S. Typhimurium* or *Shigella flexneri* (Ní Bhriain and Dorman 1993; DiNardo et al. 1982; Margolin et al. 1985). In *E. coli*, high torsional strain caused by elimination of *topA* results in transcription-driven inter-molecular RNA triplex formation, or R-loops (Drolet et al. 2003; Hraiky et al. 2000). R-loops block further transcription and stall replication forks (Higgins and Vologodskii 2004). Because *Salmonella* generates 15% lower σ , $-\sigma$ is already lower than *E. coli* strains carrying permissive *gyrB* compensatory mutations that allow introduction of *topA* amber mutations (DiNardo et al. 1982; Jaworski et al. 1991). R-loop formation would be expected to not be as severe in WT *Salmonella* as it is in *E. coli*.

Many *gyrB* mutations represented by a single nucleotide substitution in *Salmonella* are lethal when introduced into WT *E. coli*. This includes the R436S arginine to serine substitution that gives a TS phenotype at 42° in *Salmonella* but

grows normally at 30° (Pang et al. 2005). Recombination experiments that would result in exchange of the WT *Salmonella gyrB* sequence for the WT *E. coli* gene have proven to be lethal, and more surprisingly, even low level expression of either mutant or WT *S. Typhimurium* GyrB protein in WT *E. coli* is toxic (Champion and Higgins 2007). The reciprocal procedure of exchanging either WT or mutant *Salmonella gyrB* alleles with the WT *E. coli* gene is easily done, and expressing *E. coli* GyrB in *Salmonella* is non-toxic except at extremely high expression levels. The simplest hypothesis to explain these results is that the GyrB protein is an important determinant of maintaining chromosomal σ , and that the *Salmonella* GyrB protein does not generate sufficient torsional strain to meet minimum requirements for *E. coli* growth in rich medium.

seqA null mutants in *E. coli* exhibit growth-rate toxicity in rich LB medium while *Salmonella* does not (Fig. 7.7). In *E. coli* SeqA helps quell lethal consequences of over-initiation by sequestering the hemi-methylated segments of *oriC* region for 10 min or more after initiation. This inhibits DnaA-dependent replisome reassembly (Camara and Crooke 2005) and toxic problems associated with re-initiation of replication at *oriC* leads to filamentation and cell death due replication fork failure (Grigorian et al. 2003; Simmons et al. 2004). One critical step in the *oriC* initiation is unwinding of A/T rich sequences adjacent to DnaA binding sites at the origin (Baker et al. 1986; Funnell et al. 1986; 1987). The 15% lower supercoil density that *Salmonella* attains during dichotomous replication may decrease *oriC* initiation to acceptable levels in rich media without the contribution of SeqA sequestration.

An exception that proves the rule is the *mukB* mutation. Null mutants in *Salmonella* have a much more severe phenotype than *E. coli* (Fig. 7.7). An *E. coli mukB* mutant shows toxicity at 40° on LB, but grows well on LB medium at 30°. A *Salmonella mukB* mutant is viable only on minimal medium and plates on LB medium six orders of magnitude less efficiently than *E. coli*. In *E. coli*, supercoiling influences the phenotype of *mukB* mutants. Sawitzke and Austin demonstrated that growth in the presence of low levels of the GyrB inhibitor novobiocin increased the *E. coli* reliance on MukB, whereas increasing the median supercoil density by the introducing a *topA* mutation made it possible for *E. coli mukB* mutants to plate on LB up to 42° (Holmes and Cozzarelli 2000; Sawitzke and Austin 2000). *Salmonella* fits this pattern. The lower level of σ correlates with growth of *E. coli* on novobiocin and a consequential heavier reliance on MukB for condensation and segregation in rich medium (Fig. 7.7).

The fifth example is the *hns* gene. Like MukB, deletion of H-NS results in a more severe phenotype in *Salmonella* than it does in *E. coli*, where growth rates and general physiology of WT and *hns* null mutations are nearly indistinguishable (Falconi et al. 1991; Owen-Hughes et al. 1992). H-NS participates in chromosome condensation, regulation of gene expression, and the targeting of transposons (Johnson et al. 2005). In *E. coli* H-NS influences expression directly or indirectly of about 200 genes (Hommals et al. 2001) while experiments in *Salmonella* indicate that expression of over 400 genes are altered in an *hns* null mutant (Navarre et al. 2006). Most of the genes in both organisms are derepressed when H-NS is eliminated. In addition to

regulating genes, H-NS is one of the proteins that may organize loops in the bacterial chromosome (Deng et al. 2005). H-NS-induced looping in vitro has been demonstrated by atomic force microscopy (Dame et al. 2000) as well as by single molecule manipulation using optical tweezers (Dame et al. 2006). It has been proposed, although it remains to be proven, that the apparently stochastic distribution of regions of H-NS binding along the genome reflects the organization of DNA into supercoiled domains. (Noom et al. 2007). When over-expressed, H-NS can cause a dramatic condensation of the nucleoid which leads to inhibition of global transcription and an artificial stationary phase (McGovern et al. 1994). The lower natural supercoiling level of *Salmonella* may contribute to the increased dependence of this organism on H-NS for maintaining nucleoid structure. Many of the genes that are bound by H-NS in *Salmonella* appear to be A/T rich and transferred horizontally from distantly related organisms (Navarre et al. 2006).

7.13 Major Unanswered Questions

Whereas recent experimental work has made great strides in shaping a new understanding of what happens in nucleoid development, there are new questions that are being uncovered as well. For example, what machinery channels the nascent strands to form two separate nucleoids? What forces cause the *oriC* region to be more restricted than other parts of the genome? And what about loops and evolution?

7.13.1 Why Did *E. coli* and *Salmonella* Develop Different Supercoil Set Points?

One selective pressure for reducing the value of σ may come from prophages in the *Salmonella* genome (Fig. 7.4). *E. coli* K12 harbors no prophage that can make an infectious virus in a lytic cycle whereas *S. Typhimurium* has 4 lysis-competent prophage; they are Gifsy-1, Gifsy-2, Fels-1, and Fels-2 (McClelland et al. 2001). How long *Salmonella Typhimurium* has harbored 4 prophages is uncertain, but both Gifsy-1 and Gifsy-2 contain genes that contribute to pathogenicity in infected mice (Figuroa-Bossi and Bossi 1999). Another way to pose the question is to compare the behavior of the same bacteriophage in both organisms. One ideal phage for this is bacteriophage Mu, which can infect both organisms. In *Salmonella*, bacteriophage Mu is much less prone to respond to induction stimuli and the fraction of cells surviving induction is orders of magnitude higher than in *E. coli* (Champion and Higgins 2007). Thus, reduced supercoiling might have been a mechanism to reduce the toxicity of prophage in the *Salmonella* genome.

7.13.2 How Does One Prove/Disprove Loops in DNA?

The most popular theory about domain structure is that the proteins involved in creating a domain form DNA loops in vivo (Fig. 7.1) (Cook 2003; Luijsterburg et al. 2006). Electron micrographs of nucleoids from the 1970s are one reason loops originally seemed reasonable. These images illustrated a rosette of interwound supercoiled loops with a rather amorphous central region. However, alternative explanations to looping have been discussed in a recent review (Higgins et al. 2005). Biochemical studies show that two of the NAPs, H-NS and FIS can loop DNA under some circumstances (Dame et al. 2006; Skoko et al. 2006). The MukBEF complex compacts DNA with a mechanism that appears to stimulate formation of positive chiral nodes (Petrushenko et al. 2006) and MukBEF has a non-specific DNA binding affinity that could stabilize 100–500 loops. However, genetic experiments also show that the structure of the chromosome remains intact, even when many of the potential looping elements are eliminated by mutation. Must all domainins organize DNA loops that are cross-linked by proteins at the loop base (Fig. 7.1)?

Currently, no solid evidence exists for long range (10 kb) DNA loops in vivo. There are two problems that make testing loops experimentally difficult. First, stochastic boundary elements are very likely created by multiple mechanisms, which makes identifying specific candidates difficult and proving a looping mechanism in vivo extremely difficult. One could carry out experiments comparable to the chromosome conformation capture technique that shows eukaryotic chromatin loops (Dekker 2006), but what are the control predictions of a stochastic pattern? MukBEF, H-NS and FIS are three prime domainin candidates, but mutations disrupting H-NS and FIS show a only modest nucleoid phenotype in tests of supercoil structure in vivo (Hardy and Cozzarelli 2005), and cells still form nucleoids in the absence of MukBEF (Danilova et al. 2007).

Second, non-looping mechanisms clearly influence supercoil dynamics. Lynch and Wang found that two group of genes become tethered or handcuffed to the membrane by transcription coupled translation. This process restricts rotation of the DNA and can lead to hypersupercoiling of plasmids (Lynch and Wang 1993). One class of genes includes *tetA*, *lacY*, and *phoA*, which are integral membrane proteins that are inserted into the membrane via the SecA pathway. A second class of genes that become membrane attached include *tolC* and *ampC*, which are proteins exported through the cytoplasmic membrane to the outer membrane. These genes are not a minor genome fraction. Of all protein-encoding genes in *E. coli* and *Salmonella*, 25% are membrane proteins and 10% are proteins that become exported through the cytoplasmic membrane to the periplasm and outer membrane (Baars et al. 2008; Daley et al. 2005; Rey et al. 2005). Each may be a transient or stable barrier to supercoil diffusion depending on transcription/translation kinetics. Although many of these are not highly transcribed, some are and in toto they represent 1,500 sites in distributed throughout the genome that can restrict supercoil diffusion as a tether between DNA and the membrane.

7.13.3 *Are Other Nucleoid Organizing Proteins Barriers to Horizontal Gene Transfer?*

GyrB is a strong barrier to horizontal gene transfer from *Salmonella* to *E. coli*. Even relatively minor expression of the *Salmonella* GyrB in *E. coli*, which results in less than 10% of normal GyrB expression, is under strong negative selection (Champion and Higgins 2007). Moreover, introducing a single nucleotide substitution encoded by the *Salmonella* *gyrB652* mutation, or swapping the entire *E. coli* *gyrB* allele with the WT sequence from *Salmonella* is lethal. Thus, *gyrB* and all linked genes (including the *oriC* initiation region) should be excluded from recombination with *E. coli* genome, even when nucleotide restriction barriers are suppressed. Oddly, the reciprocal experiment of replacing the *Salmonella* *gyrB* locus with the *E. coli* sequence is easy to carry out. Recent evidence from gaps in sequence assembly of shotgun bacterial plasmid cloning projects shows that a surprising group of genes can act as barriers to horizontal transfer. The genes include the replication initiators, some NAPs, tRNA synthetases, outer membrane proteins, and many ribosomal proteins from close relatives of *E. coli* (Sorek et al. 2007). Some, but not all, of these inhibitory effects are caused by multi-copy expression problems. However, the only full-length genes tested for creating problems in the sequencing projects are small due to the size restrictions of 1 kb for shotgun cloning strategies. How many of the large proteins are intra-species restricted at low copy is a new biochemical and evolutionary question to think about.

Finally the comparison of segregation mechanics of *E. coli* and *Salmonella* demonstrates that intricate cooperation between different types of enzymes is necessary to complete DNA untangling of converging replication forks. DNA gyrase, Topo IV, FtsK, MukBEF, XerCD, and the *dif* site and KOPS sites are elements we know about so far. The lessons from bacteria will surely apply to the eukaryotic world because untangling DNA strands at points of replication fork convergence is every chromosome's problem. It seems improbable that a eukaryotic untangling process, which must occur at thousands of sites around the genome, will involve fewer coordinating components than are used to untangle the single chromosomes of most bacteria.

Acknowledgement Work in the Higgins Laboratory is supported by grants from the NSF – MCB9122048 and NIH – RO1GM33143.

References

- Adams DE, Shekhtman EM, Zechiedrich EL, Schmid MB, Cozzarelli NR (1992) The role of topoisomerase IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication. *Cell* 71:277–288
- Ali Azam TA, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181:6361–6370

- Baars L, Wagner S, Wickstrom D, Klepsch M, Ytterberg AJ, van Wijk KJ, de Gier JW (2008) Effects of SecE depletion on the inner and outer membrane proteomes of *Escherichia coli*. *J Bacteriol* 190:3505–3525
- Baker TA, Sekimizu K, Funnell BE, Kornberg A (1986) Extensive unwinding of the plasmid template during staged enzymatic initiation of DNA replication from the origin of the *Escherichia coli* chromosome. *Cell* 45:53–64
- Bates D, Kleckner N (2005) Chromosome and replisome dynamics in *E. coli*: Loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. *Cell* 121:899–911
- Bates D, Epstein J, Boye E, Fahrner K, Berg H, Kleckner N (2005) The *Escherichia coli* baby cell column: a novel cell synchronization method provides new insight into the bacterial cell cycle. *Mol Microbiol* 57:380–391
- Bidnenko V, Ehrlich SD, Michel B (2002) Replication fork collapse at replication terminator sequences. *EMBO J* 21:3898–3907
- Bigot S, Sivanathan V, Possoz C, Barre FX, Cornet F (2007) FtsK, a literate chromosome segregation machine. *Mol Microbiol* 64:1434–1441
- Bliska JB, Benjamin HW, Cozzarelli NR (1991) Mechanism of Tn3 resolvase recombination in vivo. *J Biol Chem* 266:2041–2047
- Bremer H, Dennis P (1996) Modulation of chemical composition and other parameters of the cell by growth rate. In: Neidhardt FC (ed) *Escherichia coli and Salmonella typhimurium*. American Society for Microbiology, Washington, DC, pp 1553
- Camara J, Crooke E (2005) Initiation of chromosomal replication. In: Higgins NP (ed) *The bacterial chromosome*. ASM, Washington, DC, pp 177–191
- Campbell JL, Kleckner N (1990) *E. coli oriC* and the *dnaA* gene promoter are sequestered from Dam methyltransferase following the passage of the chromosomal replication fork. *Cell* 62:967–979
- Castang S, McManus HR, Turner KH, Dove SL (2008) H-NS family members function coordinately in an opportunistic pathogen. *Proc Natl Acad Sci USA* 105:18947–18952
- Champion K, Higgins NP (2007) Growth rate toxicity phenotypes and homeostatic supercoil control differentiate *Escherichia coli* from *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 189:5839–5849
- Cook PR (2003) Nongenomic transcription, gene regulation and action at a distance. *J Cell Sci* 116:4483–4491
- Cortes-Ledesma F, de Piccoli G, Haber JE, Aragon L, Aguilera A (2007) SMC proteins, new players in the maintenance of genomic stability. *Cell Cycle* 6:914–918
- Cozzarelli NR, Wang JC (1990) DNA topology and its biological effects. Cold Spring Harbor Press, Cold Spring Harbor, NY
- Craigie R, Mizuuchi K (1985) Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. *Cell* 41:867–876
- Daley DO, Rapp M, Granseth E, Melen K, Drew D, von Heijne G (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* 308:1321–1323
- Dame RT, Wyman C, Goosen N (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* 28:3504–3510
- Dame RT, Noom MC, Wuite GJ (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* 444:387–390
- Danilova O, Reyes-Lamothe R, Pinskaya M, Sherratt D, Possoz C (2007) MukB colocalizes with the *oriC* region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol Microbiol* 65:1485–1492
- Dekker J (2006) The three 'C' s of chromosome conformation capture: controls, controls, controls. *Nat Methods* 3:17–21
- Deng S, Stein RA, Higgins NP (2004) Transcription-induced barriers to supercoil diffusion in the *Salmonella typhimurium* chromosome. *Proc Natl Acad Sci USA* 101:3398–3403
- Deng S, Stein RA, Higgins NP (2005) Organization of supercoil domains and their reorganization by transcription. *Mol Microbiol* 57:1511–1521

- DiNardo S, Voelkel KA, Sternglanz R, Reynolds AE, Wright A (1982) *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations in DNA gyrase genes. *Cell* 31:43–51
- Doyle M, Fookes M, Ivens A, Mangan MW, Wain J, Dorman CJ (2007) An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science* 315:251–252
- Drolet M, Broccoli S, Rallu F, Hraiky C, Fortin C, Masse E, Baaklini I (2003) The problem of hypernegative supercoiling and R-loop formation in transcription. *Front Biosci* 8:D210–D2 21
- Espeli O, Mariani KJ (2004) Untangling intracellular DNA topology. *Mol Microbiol* 52:925–931
- Espeli O, Lee C, Mariani KJ (2003) A physical and functional interaction between *Escherichia coli* FtsK and topoisomerase IV. *J Biol Chem* 278:44639–44644
- Falconi M, McGovern V, Gualerzi C, Hillyard D, Higgins NP (1991) Mutations altering chromosomal protein H-NS induce mini-Mu transposition. *New Biologist* 3:615–625
- Figuerola-Bossi N, Bossi L (1999) Inducible prophages contribute to *Salmonella* virulence in mice. *Mol Microbiol* 33:167–176
- French SL, Miller OL (1989) Transcription mapping of the *Escherichia coli* chromosome by electron microscopy. *J Bacteriol* 171:4207–4216
- Freter R, Brickner H, Fekete J, Vickerman MM, Carey KE (1983) Survival and implantation of *Escherichia coli* in the intestinal tract. *Infect Immun* 39:686–703
- Funnell BE, Baker TA, Kornberg A (1986) Complete enzymatic replication of plasmids containing the origin of the *Escherichia coli* chromosome. *J Biol Chem* 261:5616–5624
- Funnell BE, Baker TA, Kornberg A (1987) In vitro assembly of a pre-priming complex at the origin of the *Escherichia coli* chromosome. *J Biol Chem* 262:10327–10334
- Gamper HB, Hearst JE (1982) A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary, initiation and ternary complexes. *Cell* 29:81–90
- Garcia-Russell N, Harmon TG, Le TQ, Amaladas NH, Mathewson RD, Segall AM (2004) Unequal access of chromosomal regions to each other in *Salmonella*: probing chromosome structure with phage λ integrase-mediated long-range rearrangements. *Mol Microbiol* 52:329–344
- Gellert M, Mizuuchi K, O'Dea MH, Nash HA (1976) DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc Natl Acad Sci USA* 73:3872–3876
- Geszvain K, Landick R (2005) The structure of bacterial RNA polymerase. In: Higgins NP (ed) *The bacterial chromosome*. ASM, Washington, DC, pp 283–296
- Grainge I, Bregu M, Vazquez M, Sivanathan V, Ip SC, Sherratt DJ (2007) Unlinking chromosome catenanes in vivo by site-specific recombination. *EMBO J* 26:4228–4238
- Grigorian AV, Lustig RB, Guzman EC, Mahaffy JM, Zyskind JW (2003) *Escherichia coli* cells with increased levels of DnaA and deficient in recombinational repair have decreased viability. *J Bacteriol* 185:630–644
- Hardy C, Cozzarelli NR (2005) A genetic selection for supercoiling mutants of *Escherichia coli* reveals proteins implicated in chromosome structure. *Mol Microbiol* 57:1636–1652
- Hiasa H, Mariani KJ (1994) Topoisomerase III, but not topoisomerase I, can support nascent chain elongation during theta-type DNA replication. *J Biol Chem* 269:32655–32659
- Hiasa H, Mariani KJ (1996) Two distinct modes of strand unlinking during q-type DNA replication. *J Biol Chem* 271:21529–21535
- Higgins NP, Cozzarelli NR (1982) The binding of gyrase to DNA: analysis by retention to nitrocellulose filters. *Nucleic Acids Res* 10:6833–6847
- Higgins NP, Vologodskii A (2004) Topological behavior of plasmid DNA. In: Funnell BE, Phillips GJ (eds) *Plasmid biology*. ASM, Washington, DC, pp 181–201
- Higgins NP, Yang X, Fu Q, Roth JR (1996) Surveying a supercoil domain by using the gd resolution system in *Salmonella typhimurium*. *J Bacteriol* 178:2825–2835
- Higgins NP, Deng S, Pang Z, Stein R, Champion K, Manna D (2005) Domain behavior and supercoil dynamics in bacterial chromosomes. In: Higgins NP (ed) *The bacterial chromosome*. ASM, Washington, DC, pp 133–153
- Hillyard D, Edlund M, Hughes K, Marsh M, Higgins NP (1990) Subunit-specific phenotypes of *Salmonella typhimurium* HU mutants. *J Bacteriol* 172:5402–5407

- Holmes VF, Cozzarelli NR (2000) Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc Natl Acad Sci USA* 97: 1322–1324
- Hommais F, Krin E, Laurent-Winter C, Soutourina O, Malpertuy A, Le Caer JP, Danchin A, Bertin P (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* 40:20–36
- Hraiky C, Raymond M-A, Drolet M (2000) RNase H overproduction corrects a defect at the level of transcription elongation during rRNA synthesis in the absence of DNA topoisomerase I in *Escherichia coli*. *J Biol Chem* 275:11257–11263
- Huisman O, Faalen M, Girard D, Jaffé A, Toussaint A, Rouvière-Yaniv J (1989) Multiple defects in *Escherichia coli* mutants lacking HU protein. *J Bacteriol* 171:3704–3712
- Jaworski A, Higgins NP, Wells RD, Zacharias W (1991) Topoisomerase mutants and physiological conditions control supercoiling and Z-DNA formation in vivo. *J Biol Chem* 266:2576–2581
- Jeong KS, Ahn J, Khodursky AB (2004) Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol* 5:R86
- Johnson RC, Simon MI (1985) Hin-mediated site-specific recombination requires two 26 bp recombination sites and a 60 bp recombinational enhancer. *Cell* 41:781–791
- Johnson RC, Johnson LM, Schmidt JW, Gardner JF (2005) The major nucleoid proteins in the structure and function of the *E. coli* chromosome. In: Higgins NP (ed) *The bacterial chromosome*. ASM, Washington, DC, pp 65–132
- Khodursky AB, Peter BJ, Schmid MB, DeRisi J, Botstein D, Brown PO, Cozzarelli NR (2000) Analysis of topoisomerase function in bacterial replication fork movement: use of DNA microarrays. *Proc Natl Acad Sci USA* 97:9419–9424
- Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A, Darst SA (2000) A structural model of transcription elongation. *Science* 289:619–625
- Kuempel P, Henson J, Dircks L, Tecklenburg M, Lim D (1991) *dif*, a *recA*-independent recombination site in the terminus region of the chromosome of *Escherichia coli*. *New Biologist* 3:799–811
- Lesterlin C, Barre FX, Cornet F (2004) Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Mol Microbiol* 54:1151–1160
- Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci USA* 84:7024–7027
- Long AD, Mangalam HJ, Chan BY, Toller L, Hatfield GW, Baldi P (2001) Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in *Escherichia coli* K12. *J Biol Chem* 276:19937–19944
- Louarn JM, Kuempel P, Cronet F (2005) The terminus region of the *E. coli* chromosome or all's well that ends well. In: Higgins NP (ed) *The bacterial chromosome*. ASM, Washington, DC, pp 251–273
- Luijsterburg MS, Noom MC, Wuite GJ, Dame RT (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol* 156:262–272
- Lynch AS, Wang JC (1993) Anchoring of DNA to the bacterial cytoplasmic membrane through cotranscriptional synthesis of polypeptides encoding membrane proteins or proteins for export: a mechanism of plasmid hypernegative supercoiling in mutants deficient in DNA topoisomerase I. *J Bacteriol* 175:1645–1655
- Manna D, Higgins NP (1999) Phage Mu transposition immunity reflects supercoil domain structure of the chromosome. *Mol Microbiol* 32:595–606
- Manna D, Breier AM, Higgins NP (2004) Microarray analysis of transposition targets in *Escherichia coli*: The impact of transcription. *Proc Natl Acad Sci USA* 101:9780–9785
- Manna D, Porwollik S, McClelland M, Tan R, Higgins NP (2007) Microarray analysis of Mu transposition in *Salmonella enterica* serovar Typhimurium: transposon exclusion by high-density DNA binding proteins. *Mol Microbiol* 66:315–328

- Margolin P, Zumstein L, Sternglanz R, Wang JC (1985) The *Escherichia coli* *supX* locus is *topA*, the structural gene for DNA topoisomerase I. *Proc Natl Acad Sci USA* 82: 5437–5441
- McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea LWM, Stoneking T, Nhan M, Waterston R, Wilson RK (2001) The complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413:852–856
- McGovern V, Higgins NP, Chiz S, Jaworski A (1994) H-NS over-expression induces an artificial stationary phase by silencing global transcription. *Biochimie* 76:1030–1040
- Menzel R, Gellert M (1983) Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling. *Cell* 34:105–113
- Mercier R, Petit MA, Schbath S, Robin S, El Karoui M, Boccard F, Espeli O (2008) The *MatP/matS* site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* 135:475–485
- Minsky A, Kolter R (2005) Stationary phase chromosomes. In: Higgins NP (ed) *The bacterial chromosome*. ASM, Washington, DC, pp 155–166
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313:236–238
- Ní Bhriain N, Dorman CJ (1993) Isolation and characterization of a *topA* mutant of *Shigella flexneri*. *Mol Microbiol* 7:351–358
- Nielsen HJ, Ottesen JR, Youngren B, Austin SJ, Hansen FG (2006) The *Escherichia coli* chromosome is organized with the left and right chromosome arms in separate cell halves. *Mol Microbiol* 62:331–338
- Niki H, Jaffe A, Imamura R, Ogura T, Hiraga S (1991) The new gene *mukB* codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of *E. coli*. *EMBO J* 10:183–193
- Niki H, Yamaichi Y, Hiraga S (2000) Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14:212–223
- Noom MC, Navarre WW, Oshima T, Wuite GJ, Dame RT (2007) H-NS promotes looped domain formation in the bacterial chromosome. *Curr Biol* 17:R913–914
- O'Donnell M (2006) Replisome architecture and dynamics in *Escherichia coli*. *J Biol Chem* 281:10653–10656
- Oram M, Travers AA, Howells AJ, Maxwell A, Pato ML (2006) Dissection of the bacteriophage Mu strong gyrase site (SGS): significance of the SGS right arm in Mu biology and DNA gyrase mechanism. *J Bacteriol* 188:619–632
- Owen-Hughes TA, Pavitt GD, Santos DS, Sidebotham JM, Hulton CSJ, Hinton JCD, Higgins CF (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* 71:255–266
- Pang Z, Chen R, Manna D, Higgins NP (2005) A gyrase mutant with low activity disrupts supercoiling at the replication terminus. *J Bacteriol* 187:7773–7783
- Perederina A, Svetlov V, Vassilyeva MN, Tahirov TH, Yokoyama S, Artsimovitch I, Vassilyev DG (2004) Regulation through the secondary channel–structural framework for ppGpp-DksA synergism during transcription. *Cell* 118:297–309
- Peter BJ, Ullsperger C, Hiasa H, Marians KJ, Cozzarelli NR (1998) The structure of supercoiled intermediates in DNA replication. *Cell* 94:819–827
- Peter BJ, Arsuaga J, Breier AM, Khodursky AB, Brown PO, Cozzarelli NR (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol* 5:R87
- Petrushenko ZM, Lai CH, Rai R, Rybenkov VV (2006) DNA reshaping by MukB. Right-handed knotting, left-handed supercoiling. *J Biol Chem* 281:4606–4615
- Pettijohn DE (1996) The nucleoid. In: Neidhardt FC (ed) *Escherichia coli and Salmonella*. Cellular and Molecular Biology, vol 1, 2nd edn. ASM, Washington, DC, pp 158–166

- Pettijohn DE, Pfenninger O (1980) Supercoils in prokaryotic DNA restrained in vivo. *Proc Natl Acad Sci USA* 77:1331–1335
- Postow L, Crisona NJ, Peter BJ, Hardy CD, Cozzarelli NR (2001) Topological challenges to DNA replication: conformations at the fork. *Proc Natl Acad Sci USA* 98:8219–8226
- Pruss GJ, Franco RJ, Chevalier SG, Manes SH, Drlica K (1986) Effects of DNA gyrase inhibitors in *Escherichia coli* topoisomerase I mutants. *J Bacteriol* 168:276–282
- Rahmouni AR, Wells RD (1992) Direct evidence for the effect of transcription on local DNA supercoiling in vivo. *J Mol Biol* 223:131–144
- Raji A, Zabel DJ, Laufer CS, Depew RE (1985) Genetic analysis of mutations that compensate for loss of *Escherichia coli* DNA topoisomerase I. *J Bacteriol* 162:1173–1179
- Rey S, Acab M, Gardy JL, Laird MR, deFays K, Lambert C, Brinkman FS (2005) PSORTdb: a protein subcellular localization database for bacteria. *Nucleic Acids Res* 33:D164–D168
- Rybenkov V, Ullsperger C, Vologodskii A, Cozzarelli NR (1997) Simplification of DNA topology below equilibrium values by type II topoisomerases. *Science* 277:690–693
- Saldanha R, Flanagan P, Fennwald M (1987) Recombination by resolvase is inhibited by *lac* repressor simultaneously binding operators between *res* sites. *J Mol Biol* 196:505–516
- Sawitzke JA, Austin S (2000) Suppression of chromosome segregation defects of *Escherichia coli* *muk* mutants by mutations in topoisomerase I. *Proc Natl Acad Sci USA* 97:1671–1676
- Scheirer K, Higgins NP (2001) Transcription induces a supercoil domain barrier in bacteriophage Mu. *Biochimie* 83:155–159
- Schmid MB, Johnson RC (1991) Southern revival-news of bacterial chromatin. *New Biologist* 3:945–950
- Simmons LA, Breier AM, Cozzarelli NR, Kaguni JM (2004) Hyperinitiation of DNA replication in *Escherichia coli* leads to replication fork collapse and inviability. *Mol Microbiol* 51:349–358
- Sinden RR, Pettijohn DE (1981) Chromosomes in living *Escherichia coli* cells are segregated into domains of supercoiling. *Proc Natl Acad Sci USA* 78:224–228
- Sinden RR, Carlson JO, Pettijohn DE (1980) Torsional tension in the DNA double helix measured with trimethylpsoralen in living *E. coli* cells: analogous measurements in insect and human cells. *Cell* 21:773–783
- Skarstad K, Torheim N, Wold S, Lurz R, Messer W, Fossum S, Bach T (2001) The *Escherichia coli* SeqA protein binds specifically to two sites in fully and hemimethylated *oriC* and has the capacity to inhibit DNA replication and affect chromosome topology. *Biochimie* 83:49–51
- Skoko D, Yoo D, Bai H, Schnurr B, Yan J, McLeod SM, Marko JF, Johnson RC (2006) Mechanism of chromosome compaction and looping by the *Escherichia coli* nucleoid protein fis. *J Mol Biol* 364:777–798
- Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EM (2007) Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* 318:1449–1452
- Staczek P, Higgins NP (1998) DNA gyrase and Topoisomerase IV modulate chromosome domain size in vivo. *Mol Microbiol* 29:1435–1448
- Stein R, Deng S, Higgins NP (2005) Measuring chromosome dynamics on different timescales using resolvases with varying half-lives. *Mol Microbiol* 56:1049–1061
- Stone MD, Bryant Z, Crisona NJ, Smith SB, Vologodskii A, Bustamante C, Cozzarelli NR (2003) Chirality sensing by *Escherichia coli* topoisomerase IV and the mechanism of type II topoisomerases. *Proc Natl Acad Sci USA* 100:8654–8659
- Sueoka N (1971) The chromosome of *Bacillus subtilis*. II: A general solution for the age distribution function of bacterial chromosomes. *Genetics* 68:349–358
- Sunako Y, Onogi T, Hiraga S (2001) Sister chromosome cohesion of *Escherichia coli*. *Mol Microbiol* 42:1233–1241
- Tupper AE, Owen-Hughes TA, Ussery DW, Santos DS, Ferguson DJP, Sidebotham JM, Hinton JCD, Higgins CF (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J* 13:258–268
- Valens M, Penaud S, Rossignol M, Cornet F, Boccard F (2004) Macrodome organization of the *Escherichia coli* chromosome. *EMBO J* 23:4330–4341

- Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Natl Acad Sci USA* 101:9257–9262
- Vologodskii AV, Zhang W, Rybenkov VV, Podtelezhnikov AA, Subramanian D, Griffith JD, Cozzarelli NR (2001) Mechanism of topology simplification by type II DNA topoisomerases. *Proc Natl Acad Sci USA* 98:3045–3049
- Wang JC (1971) Interaction between DNA and an *Escherichia coli* protein omega. *J Mol Biol* 55:523–533
- Wang JC (1991) DNA topoisomerases: why so many? *J Biol Chem* 266:6659–6662
- Wang X, Possoz C, Sherratt DJ (2005) Dancing around the divisome: asymmetric chromosome segregation in *Escherichia coli*. *Genes Dev* 19:2367–2377
- Wang X, Liu X, Possoz C, Sherratt DJ (2006) The two *Escherichia coli* chromosome arms locate to separate cell halves. *Genes Dev* 20:1727–1731
- Wang X, Reyes-Lamothe R, Sherratt DJ (2008) Modulation of *Escherichia coli* sister chromosome cohesion by topoisomerase IV. *Genes Dev* 22:2426–2433
- Weitao T, Dasgupta S, Nordstrom K (2000) Role of the *mukB* gene in chromosome and plasmid partition in *Escherichia coli*. *Mol Microbiol* 38:392–400
- Willmott CJR, Critchlow SE, Eperon IC, Maxwell A (1994) The complex of DNA gyrase and quinolone drugs with DNA forms a barrier to transcription by RNA polymerase. *J Mol Biol* 242:351–363
- Woo JS, Lim JH, Shin HC, Suh MK, Ku B, Lee KH, Joo K, Robinson H, Lee J, Park SY, Ha NC, Oh BH (2009) Structural studies of a bacterial condensin complex reveal ATP-dependent disruption of intersubunit interactions. *Cell* 136:85–96
- Yin YW, Steitz TA A (2002) Structural basis for the transition from Initiation to elongation transcription in T7 RNA polymerase. *Science* 298:1387–1395
- Zechiedrich EL, Khodursky AB, Bachellier S, Schneider R, Chen D, Lilley DM, Cozzarelli NR (2000) Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli*. *J Biol Chem* 275:8103–8113
- Zhang A, Belfort M (1992) Nucleotide sequence of a newly-identified *Escherichia coli* gene, *stpA*, encoding an H-NS-like protein. *Nucleic Acids Res* 20:6735
- Booker BM, Deng S, Higgins NP (Submitted) Topological structure of highly transcribed operons in *Salmonella typhimurium*

Chapter 8

Nucleoid-Associated Proteins: Structural Properties

Ümit Pul and Rolf Wagner

Abstract This chapter summarizes the structural details and functional properties of a representative selection of nucleoid-associated proteins (NAPs), consisting of FIS, H-NS, LRP, IHF, HU and Dps. Currently, high resolution structural information of the complete proteins is available for FIS, LRP, IHF, HU and Dps, while for H-NS structures have only been solved for the separate C- and N-terminal halves of the protein. In some cases, such as IHF or HU, structures of the proteins in complex with their target DNA have been determined. For the other proteins reasonable models describing the architecture of the respective protein-DNA complexes have been derived from biochemical and biophysical studies.

A common denominator for all the nucleoid-associated proteins is their property to induce structural deformations to the bound DNA. Each protein appears to affect DNA conformation or topology in a very specific way, however. Despite their general function in compacting DNA these differential properties are key to the high versatility of the NAPs as general and gene-specific regulators in the cell.

Keywords Bacterial nucleoid • crystal structure • DNA topology • protein-DNA interaction • transcription factor

8.1 Introduction

The dynamic structure of bacterial genomes consists of negatively supercoiled DNA loops to which a family of small DNA-binding proteins is associated. These proteins are known as nucleoid-associated proteins (NAPs). They are involved in modulating the dynamic structure of the chromosome and fulfil important tasks in

Ü. Pul and R. Wagner (✉)

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1,
D-40225, Düsseldorf

e-mail: R.Wagner@rz.uni-duesseldorf.de

replication, recombination and several members of the NAP family also act as direct regulators of transcription. Since about half of the negative supercoils in bacterial genomes are constrained by these proteins and negative DNA supercoiling has itself an important regulatory role in bacteria, NAPs also affect the gene expression pattern of the cell through changing the local supercoiling of the DNA.

The individual members of the NAP family have in common that they are DNA-binding proteins with modular domain composition. The specificity of binding to the target DNA ranges from sequence-specific binding over DNA conformation-specific binding to non-sequence binding. For most but not all sequence-specific binders typical DNA recognition motifs, such as helix-turn-helix or β -sheet motifs, are responsible, while for some other NAPs the molecular mechanisms of DNA interaction are still unclear. In this chapter the structural properties important to understand the functional versatility of this group of proteins will be described in detail for some representative members of the family.

8.2 FIS: The Factor for Inversion Stimulation

FIS is a homodimeric protein of 22 kDa found as a global regulator in enteric bacteria. One of the most pronounced characteristics of FIS is its growth-phase dependent expression. In early logarithmic growth FIS is the most abundant nucleoid-associated protein, whereas during entry into stationary phase this number drops to about 100 copies per cell.

The structure of FIS has been solved by X-ray crystallography (Kostrewa et al. 1991; Kostrewa et al. 1992). As can be seen from the structure shown in Fig. 8.1 FIS is composed of two intertwined α -helical subunits, forming a globular structure with four bundled α -helices (A, B, C, D). The C-terminal helices (helices CD) constitute the helix-turn-helix (HTH) DNA binding motif. Recognition occurs via helix D, which fits into the major groove of the target DNA. FIS binding sites usually contain a strongly degenerated palindromic recognition motif with the consensus sequence (G/T)NN(C/T)(A/G)NN(A/T)NN(C/T)(A/G)NN(C/A) (Pan et al. 1994). According to the high-resolution structure of FIS dimers the distance of the two helices D and D' is 2.4 nm (Kostrewa et al. 1992). Thus, the target DNA must be bent by FIS in order to fit the two helices into two adjacent major grooves of the DNA. FIS binding therefore leads to bending of the DNA between 50° and 90° (Pan et al. 1996). At high concentration FIS also binds DNA non-specifically (Schneider et al. 2001; Skoko et al. 2006), which is of biological significance, considering its abundance during logarithmic growth. This view is supported by the finding that FIS interacts with more than 10,000 sites also present at intergenic regions within the *E. coli* genome (Cho et al. 2008). Similar observations have been made for H-NS and LRP (Grainger et al. 2006). The amino acid side-chains of FIS, most critical for recognition and binding to specific DNA sequences, have been identified by mutagenesis and are indicated in Fig. 8.1 (Feldman-Cohen et al. 2006). Moreover, nucleotide substitution experiments indicated that multiple contacts

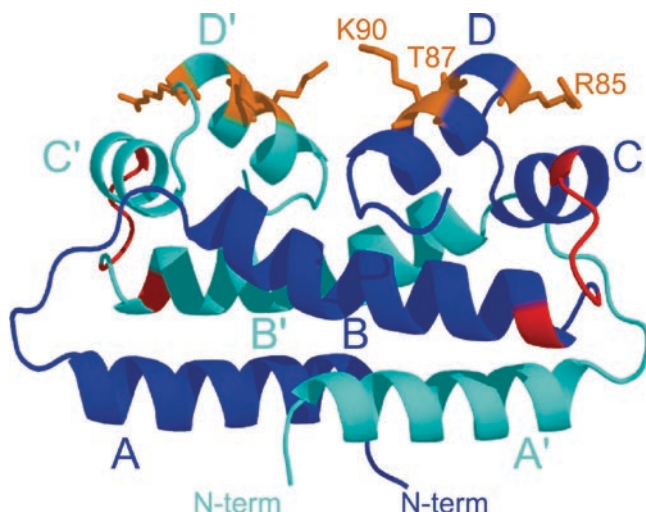


Fig. 8.1 Crystal structure of FIS dimer (PDB ID: 1ETY; (Cheng et al. 2000)). The active form of FIS is a homodimer, which binds to the major groove of a degenerated palindromic DNA sequence via its helix-turn-helix motif located in the C-terminal domain. α -helices for both monomers (A/A' to D/D') are indicated. Amino acids K90, T87 and R85, critical for recognition of specific DNA sequences are depicted in orange (Feldman-Cohen et al. 2006). The region responsible for RNA polymerase recruitment (activation patch) is coloured in red (Bokal et al. 1997; Cheng et al. 2000)

between FIS and the DNA phosphate backbone with only a small number of base contacts contribute to the specificity of these interactions (Shao et al. 2008a, b). This observation reflects the ability of FIS to bind to poorly related sequences with relatively high affinity. The involvement of many phosphate contacts could also explain the positive correlation between FIS binding and the high AT-content of the target sites, as observed by genome-wide ChIP-chip analyses (Grainger et al. 2006; Shao et al. 2008a).

The most outstanding physiological function of FIS certainly consists in its potential to activate stable RNA (tRNA and rRNA) transcription (Nilsson et al. 1990, 1992). FIS binds to multiple binding sites within the upstream regulatory region of all ribosomal RNA operons and activates transcription (Hirvonen et al. 2001; Hillebrand et al. 2005). The activation mechanism involves recruitment of RNA polymerase to the promoter through direct interaction with the C-terminal domain of the α -subunit of RNA polymerase (Bokal et al. 1997). The FIS side-chains important for activation at ribosomal RNA promoters have been identified by mutagenesis and the respective protein domain is termed 'the activation patch' (see Fig. 8.1) (Bokal et al. 1997; Cheng et al. 2000). In addition, FIS can also influence transcription activity through introducing microloops that cause negative supercoils, which in turn facilitates promoter melting (Muskhelishvili et al. 1997). For the *rnmB* P1 promoter it was shown that another nucleoid-associated protein, namely H-NS, acts as an antagonist to FIS-mediated activation (see below) (Afflerbach et al. 1999). An alternative mechanism has been suggested by a recent study to elucidate the

regulation of Dps expression by FIS (Grainger et al. 2008). In this case, FIS traps RNA polymerase $E\sigma^{70}$ at the promoter and thereby blocks the binding of the alternative holoenzyme, $E\sigma^{38}$. At stationary phase this repression is abrogated due to decreased FIS concentrations. Thus, FIS is not only capable of interacting with DNA but also with other regulatory proteins. Hin recombinase or RNA polymerase are additional examples of such heterologous protein interactions (Pan et al. 1996). Actually, FIS-regulated genes often contain multiple binding sites, which are found upstream of promoters, where they confer activation but also overlapping or downstream of promoters, as is typical for transcriptional repression. Thus, FIS is a global regulator with a broad set of versatile regulatory mechanisms (nucleoid-structuring, RNA polymerase trapping, interference with RNA polymerase binding or through changes in DNA topology) (Chapter 14).

8.3 H-NS: Histone-Like Nucleoid Structuring Protein

H-NS has been characterized as both a DNA-structuring protein and transcriptional regulator with predominant repressing functions (Schröder and Wagner 2002; Rimsky 2004; Fang and Rimsky 2008). It is an impressive protein because of its manifold influences on diverse cellular processes, such as adaptation to altered growth or stress conditions, regulation of bacterial virulence, DNA compaction and silencing and, most interestingly, the recently identified involvement in the defence against foreign DNA (Lucchini et al. 2006; Navarre et al. 2007). H-NS was first described as a major component of the bacterial nucleoid (Varshavsky et al. 1977). It is a small DNA binding protein with a highly conserved structure and function in Gram-negative bacteria (Bertin et al. 1999). In *Escherichia coli* H-NS has a molecular weight of 15.6 kDa and neutral pI. The functional form is a dimer or larger multimer (Ceschini et al. 2000). Today it is clear that H-NS is a global regulator of the bacterial cell, which together with other members of the nucleoid-associated proteins contributes to the efficient adaptation of bacteria to different environmental conditions. Despite numerous investigations to understand the molecular mechanism of DNA binding and oligomerization of H-NS, we still do not know the exact details of these interactions.

As yet, no crystal structure of the complete H-NS protein is available. However, 3D information has been derived from NMR data of the truncated N- and C-terminal domains (Shindo et al. 1995, 1999; Renzoni et al. 2001; Bloch et al. 2003; Esposito et al. 2002). As shown in Fig. 8.2, H-NS can be divided into two structural and functional domains, the N-terminal dimerization domain (Fig. 8.2a, b) and C-terminal DNA-binding domain (Fig. 8.2c). The independently folded domains are connected by an unstructured flexible linker, which probably hinders the crystallization of the native protein. The domain composition of H-NS is also supported by genetic analyses. For example, a truncated H-NS protein from *E. coli*, comprising only the first 46 N-terminal amino acid residues, has been described as the smallest H-NS derivative to dimerize, but incapable of forming

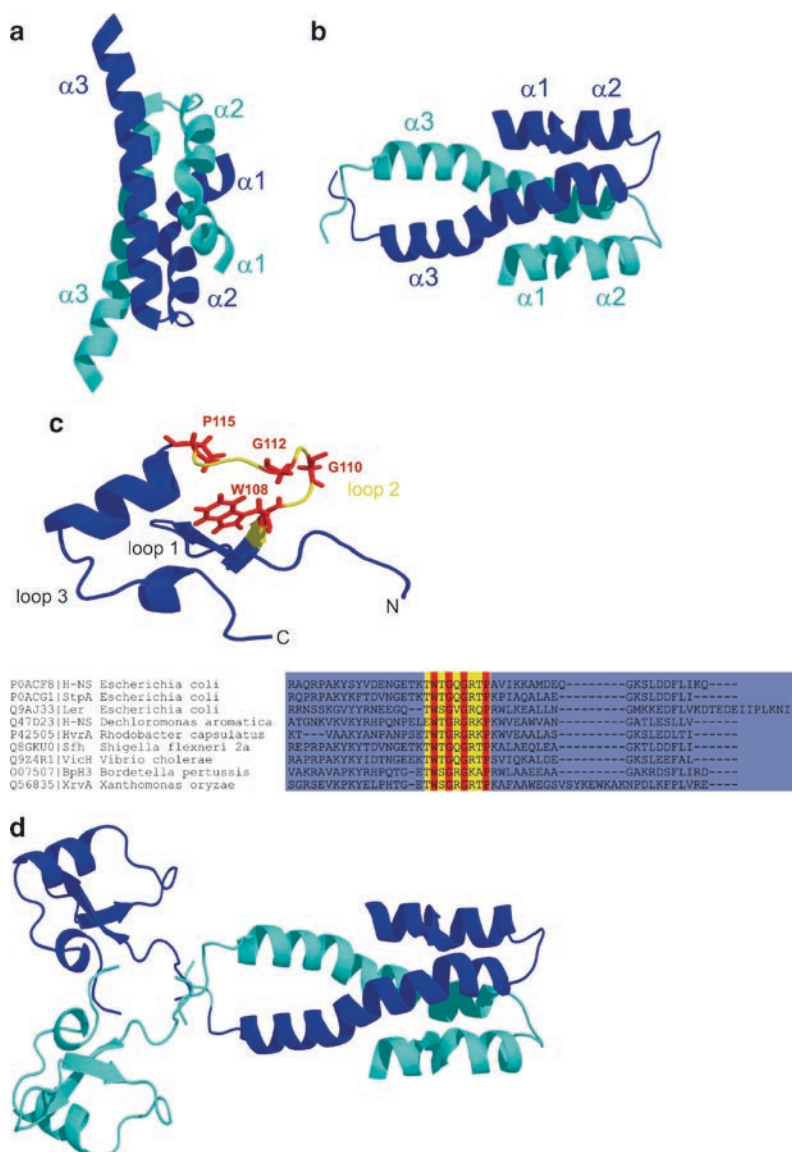


Fig. 8.2 3-dimensional structures of the N- and C-terminal domains of H-NS. **(a)** The NMR-structure of the dimerization domain (residues 1–46) of H-NS is shown (PDB ID: 1NI8; (Bloch et al. 2003)). The α -helices for both monomers are indicated ($\alpha 1$, $\alpha 2$ or $\alpha 3$). **(b)** An alternative NMR structure determined for the H-NS N-terminal domain (residues 1–57) from *S. typhimurium* is shown (PDB ID: 1LR1; (Esposito et al. 2002)). **(c)** The three-dimensional NMR structure of the C-terminal domain of H-NS (residues 91–137) is shown (PDB ID: 1HNR; (Shindo et al. 1995)). The conserved residues important for DNA binding are indicated in red. At the bottom an amino acid sequence alignment and accession numbers of the C-terminal residues from several members of the H-NS family is shown. The amino acid residues within loop 2 are indicated in yellow as in the three-dimensional structure above. Identical amino acids are coloured in red. **(d)** 3D structure model of an H-NS dimer modeled from separate structural domains as in **(b)** and **(c)**

higher oligomers (Bloch et al. 2003). The structure of the N-terminal H-NS₁₋₄₆ obtained from 2D NMR analysis is shown in Fig. 8.2a. Dimerization occurs by antiparallel coiled-coil association of the α 3-helices through hydrophobic residues. The α 1 and α 2 helices are also involved and further stabilize the dimeric structure (Bloch et al. 2003). Interestingly, according to the three-dimensional structure of the H-NS₁₋₅₇ dimer from *Salmonella typhimurium* determined in a separate study (Esposito et al. 2002) the α 3-helices are oriented in parallel, even though the overall structures of the monomers are almost identical in both studies (Fig. 8.2b). In both structures the two short helices α 1 and α 2 are folded back to the α 3-helices, but in the antiparallel orientation the α 2 helix of one H-NS monomer is associated with the α 3-helix of the other protomer and not with that of the same molecule, as in the case of the parallel orientation. The observed differences might be due to the additional amino acid residues in the H-NS₁₋₅₇ variant compared to the H-NS₁₋₄₆ with the antiparallel orientation of the long helices.

Although H-NS can clearly be divided into two structural domains, the assignment of independent functions for the individual domains (with respect to dimerization/oligomerization and DNA binding) has turned out to be difficult. Generally, mutations in the N-terminal domain can affect DNA binding, and mutations in the C-terminal domain affect the dimerization/oligomerization of H-NS. Since dimerization/oligomerization is an important prerequisite for DNA-binding, interpretations from binding studies with H-NS mutants are notoriously difficult.

Amino acid sequences potentially involved in oligomerisation of H-NS in vivo were analysed in a study using chimeric constructs in which wild-type or mutant H-NS fragments had been fused to a λ phage repressor (Stella et al. 2005). The analyses revealed that deletion of a proline residue in the C-terminal domain (Δ P115) or substitution of the same amino acid by alanine (P115A) impeded dimerization, and to a greater extent, tetramerization of the chimeric H-NS constructs in vivo (see Fig. 8.2c). In a separate investigation it was shown by gel retardation and footprint analyses that the mutant H-NS variant Δ P115 also displays reduced affinity for curved DNA fragments (Spurio et al. 1997). It should be noted that binding to curved DNA constitutes actually the decisive specificity for H-NS-DNA recognition and that it can also discriminate between short synthetic bent DNA fragments (Yamada et al. 1990). The H-NS Δ P115 mutant has lost this preferential binding, suggesting that oligomerization and DNA binding specificity may be directly related. Similar results were also reported for amino acid substitution mutants, where residues within the C-terminal domain, including the proline residue 115, had been replaced (Badaut et al. 2002). Again, these mutants can bind non-specifically to DNA but they have lost the preference for curved DNA. Together the data indicate that the C-terminal domain contains the DNA motif responsible for specific binding, but to some extent is also involved in protein-protein interaction. Moreover, it has been reported that the flexible linker separating the N- and C-terminal domains may also participate in oligomerization of H-NS. This conclusion was derived from analyses, which show that the interactions leading to the formation of H-NS tetramers are thermodynamically different compared to those leading to H-NS dimers (Ceschini et al. 2000). Thus, it has been

suggested that after dimerization of H-NS through hydrophobic interactions of the N-terminal domain, further oligomerization of the H-NS dimers involves amino acids within the linker region and also partly amino acids from the C-terminal domain. Note that there may be a difference between the solution state and the DNA-bound state; related to that the current evidence suggests that H-NS is bound in dimeric form when it bridges DNA (Dame et al. 2006). Oligomerization of H-NS is important for DNA binding and the DNA bending capacity, while dimerization is a prerequisite for the formation of tetramers or higher oligomers. Therefore it is difficult to assign oligomerization or DNA binding to one structural domain, even though the necessary individual functions may be catalyzed by individual domains of the protein (Chapter 13).

The part of the H-NS structure responsible for DNA interaction does not correspond to any conserved amino acid folds known for DNA interaction. It is clearly folded, however, in a separate domain, which has a rather large spatial distance to the N terminus of roughly 4.5 nm as revealed by FRET analysis between Trp 108 in the centre of the C-terminal binding domain and the N-terminus of the protein (Schröder et al. 2001). Details of the structure comprising the last 47 C-terminal amino residues have been solved by NMR and are presented in Fig. 8.2c (Shindo et al. 1995, 1999). According to this analysis the C-terminal domain consists of four small loop structures separated by two short β -strands and two helical segments. A series of unstructured amino acids, forming loop 2 harbours the amino acid residues, which, according to mutagenesis studies, are most likely involved in direct DNA interaction (see below). This loop is flanked by a small α -helix and a 3_{10} -helix close to the C-terminus. As also shown in Fig. 8.2c the C-terminal DNA binding domain is highly conserved in all H-NS-like proteins, suggesting similar structures of the DNA binding regions. Within that structure four amino acid residues (W108, G110, G112 and P115) are notably important. They are conserved in nearly all H-NS-like proteins currently collected in the public databases. These highly conserved amino acid residues are all located within loop 2. Based on those sequence alignments the following consensus motif for an H-NS-type DNA binding domain has been proposed: TWTG-GR-P (Dorman et al. 1999). Previous analyses using amino acid substitutions within the C-terminal domain of H-NS support the direct involvement of loop 2 in DNA binding (Ueguchi et al. 1996). A G112D mutant of H-NS, for instance, is strongly impeded in DNA binding and also does not show a comparable effect on DNA topology as the wild-type protein (Ueguchi et al. 1996; Pul et al. 2005, 2007). Despite this information the exact molecular details of H-NS-DNA interaction are still unknown. Such an analysis is primarily hampered by the fact that for long no clearly defined DNA consensus binding sequence was known. Even the question of whether H-NS binds to the major or minor groove of DNA has not been answered with confidence. A recent study (Sette et al., 2009) reports evidence for binding of the C-terminal fragment of H-NS to the minor groove of a conserved DNA element. The only fact which is undisputed about H-NS binding, concerns its ability to interact selectively with curved or flexible DNA segments. Such DNA properties generally correlate with a high content of AT base pairs, which, when clustered in helical phase, result in curved DNA.

It is not surprising, therefore, that primary H-NS binding sites have been mapped in AT-rich sequence regions. A recent proposal of such a primary ‘high affinity’ H-NS binding site (tCGATAAATT) in principle reflects this high AT content (Bouffartigues et al. 2007; Lang et al. 2007). Often the initial binding step involves such AT-rich sites and further interactions are followed by cooperative polymerization of H-NS molecules along the DNA, which ultimately results in protein coating of rather large DNA regions. This protein-coating either hinders binding of RNA polymerase to promoter DNA or it may interfere with the interaction of other regulatory proteins. The coating mechanism as a consequence of cooperative H-NS binding is supported by the fact that generally large regions of the DNA are protected in DNase I footprint experiments (Pul et al. 2005; Bouffartigues et al. 2007). The fact that the variation of the distance and angular orientation of a synthetic curved H-NS binding module does not show significant effects on transcription inhibition at different promoters is also consistent with the coating mechanism (Pul et al. 2008).

The capacity of H-NS for lateral oligomerization along the DNA and the presence of two DNA binding domains per H-NS dimer allows for the simultaneous interaction with two DNA strands, which gives rise to DNA bridging (Fig. 8.2d). Such a bridging activity of H-NS and other H-NS-like proteins has been visualized in AFM studies (Dame et al. 2000, 2005). DNA bridging may certainly contribute to DNA compaction of the bacterial nucleoid. More importantly, however, bridging of two adjacent DNA strands also brings about transcription regulation. As shown for the *rnmB* P1 promoter, H-NS-mediated bridging leads to RNA polymerase trapping in an open promoter complex, which is inadequate for transcription (Schröder and Wagner 2000; Dame et al. 2002). A trapping mechanism through DNA bridging has also been described for the *hdeAB* promoter. In this special case sigma-factor selectivity is regulated because H-NS-mediated bridging occurs only with $E\sigma^{70}$ but not with $E\sigma^{38}$ (Shin et al. 2005). Interestingly, sigma-factor selectivity is also conferred by other NAPs, such as IHF and LRP at the *osmY* promoter (Colland et al. 2000) or FIS at the *dps* promoter (Grainger et al. 2008).

Like most of the other NAP members H-NS often acts in concert either with other members of the NAP family or with other DNA binding proteins, thereby displaying synergistic or antagonistic effects. This general property of concerted function significantly adds to the regulatory variability of the NAP family of regulators. Well studied examples are the seven ribosomal RNA operons in *E. coli*, where FIS-dependent activation of transcription is counteracted by H-NS, while transcription inhibition by LRP is supported (Afflerbach et al. 1999; Pul et al. 2005). Other examples of antagonism include the virulence gene transcriptional regulator *VirB* in *Shigella flexneri* (Turner and Dorman 2007), transcription of the *cspA* mRNA, which is under the antagonistic control of FIS and H-NS (Brandi et al. 1999) or *hns* expression itself, which is under H-NS autoregulation antagonized by FIS (Falconi et al. 1996; Stoebel et al. 2008). In line with the view of a concerted function of different NAP members are the binding profiles of FIS, H-NS and IHF, which have been analyzed on a genome-wide scale. This analysis reveals considerable overlap

in binding of the three proteins and also shows that sites of FIS and H-NS binding are often close to sites of RNA polymerase association (Grainger et al. 2006).

H-NS does not only make contact with itself, forming homodimers or higher oligomers but is also known to undergo heteromeric contacts with a variety of other proteins. Notable examples are the heterodimeric forms with StpA or other H-NS homologs, which share the dimerization domain. However, heteromeric contacts are also formed with more distantly related proteins, such as FliG, Hfq, Hha or the phage T7 gene product 5.5 (Kajitani and Ishihama 1991; Liu and Richardson 1993; Williams et al. 1996; Donato and Kawula 1998; Nieto et al. 2002). Such combinations of regulatory proteins may furnish the cell with a battery of new regulatory tools, providing novel specificity and tuning mechanisms. This may be of special importance for adaptation reactions under rapidly changing environmental conditions and is consistent with the known involvement of H-NS in modulation of bacterial gene expression in response to temperature and osmolarity (Falconi et al. 1998; Nieto et al. 2002; Dorman 2004).

Ever since their early identification NAPs have been recognized as important elements for the transfer of genetic elements. This is partly reflected in their names, which in case of FIS and IHF are acronyms of their functions characterized in the first place. FIS (factor for inversion stimulation) is involved in phage recombination reactions and IHF (integration host factor) also participates in phage λ integration. For H-NS it was recently shown that it binds to genes that were acquired by horizontal gene transfer, where it functions as gene silencer. Binding to the foreign genes depends on a higher AT content relative to the resident genome, consistent with the preference of H-NS to bind to AT-rich DNA sequences. This has led to suggest another interesting new function for H-NS, namely transcriptional silencing of AT-rich foreign DNA obtained through horizontal gene transfer (Lucchini et al. 2006; Dorman 2007; Navarre et al. 2007).

In summary, the modular composition and structural properties of H-NS, unrelated to conventional transcription factors, provide this protein with a variety of unusual functions and a wide range of distinct properties, making it a highly versatile bacterial regulator to efficiently adapt the bacterial gene expression pattern to altered conditions.

8.4 LRP: Leucine Responsive Regulatory Protein

Leucine responsive regulatory proteins constitute a family of widespread transcriptional regulators with global functions in cellular adaptation to environmental changes. *E. coli* LRP is a basic protein with a pI of 9.3 and exists in solution as homodimer composed of two identical subunits with a molecular weight of 18.8 kDa each. As a transcriptional activator or repressor LRP is mainly involved in the regulation of amino acid metabolism-related genes. Hence, LRP has been assigned to the family of so-called 'feast and famine' regulatory proteins (FFRPs) found in eubacteria and archaea, which contribute to the adaptation of

the bacterial gene expression in response to the nutritional quality (Yokoyama et al. 2006).

The cellular concentration of LRP strongly depends on the availability of nutrients and is inversely proportional to the bacterial growth rate (Landgraf et al. 1996). DNA microarray analysis has shown that at least 10% of *E. coli* genes are affected by LRP; most of these genes are involved in the adaptation of *E. coli* to nutrient limitations (Tani et al. 2002).

The oligomerization state of many FFRPs can be regulated through allosteric effector molecules. In the case of *E. coli* LRP the amino acid leucine is the allosteric effector. This is a unique feature among the NAP family members for which otherwise no allosteric effector molecules are known. Through allosteric modulation of the quarternary structure leucine influences the association state of *E. coli* LRP, which results in altered DNA binding affinity (Chen and Calvo 2002). The resulting regulatory effects for individual promoters differ, however. Hence, among the LRP-regulated genes leucine can either enhance, reduce or have no effect on binding of LRP (Calvo and Matthews 1994). The association reactions of LRP are complex and it is assumed that at nanomolar concentrations LRP dimers are the predominant species. Determination of the LRP-DNA stoichiometry indicated that LRP can bind as a dimer but higher aggregates are frequently found in DNA complexes (Chen et al. 2005). At micromolar concentrations LRP exists as octamers and hexadecamers, which in the presence of leucine dissociate preferentially to octamers (Chen et al. 2001; Chen and Calvo 2002). The exact mechanism behind the differential effect of leucine on LRP binding is not fully understood, however.

The three-dimensional structure of *E. coli* LRP has been solved by X-ray crystallography (de los Rios and Perona 2007). Figure 8.3a shows the modular structure of an LRP monomer, which is characterized by two structural domains also reflecting the different functions of LRP. The N-terminal domain contains the helix-turn-helix DNA binding motif, consisting of three α -helices. According to mutational analysis the $\alpha 3$ -helix of this motif was found to be responsible for binding to the major groove of the DNA (Platko and Calvo 1993). The C-terminal domain of LRP contains two α -helices flanking four antiparallel β -strands, forming a $\alpha\beta$ -sandwich. The resulting $\beta\alpha\beta\alpha\beta$ -structure constitutes a structural motif found in many LRP-like regulatory proteins or enzymes involved in amino acid metabolism and has been termed RAM domain (regulation of amino acid metabolism) (Ettema et al. 2002). Within the RAM domain the $\beta 3$ – $\beta 4$ loop contains the binding site for the effector amino acid leucine. The N- and C-terminal domains are linked by a flexible linker, containing a β -strand (depicted as ‘crossover β -strand’ in Fig. 8.3a according to de los Rios and Perona (2007)), which upon dimerization forms a two-stranded antiparallel β -sheet with the corresponding β -strand of the second monomer at the homodimer centre (Fig. 8.3b). In addition, dimerization is attributed to hydrophobic interactions between the two β -sheets of the RAM domains (Chapter 15).

As shown in Fig. 8.3c four LRP dimers associate to form an octamer, arranged as a disc, in a similar way as has been described for LrpA from *Pyrococcus furi-*

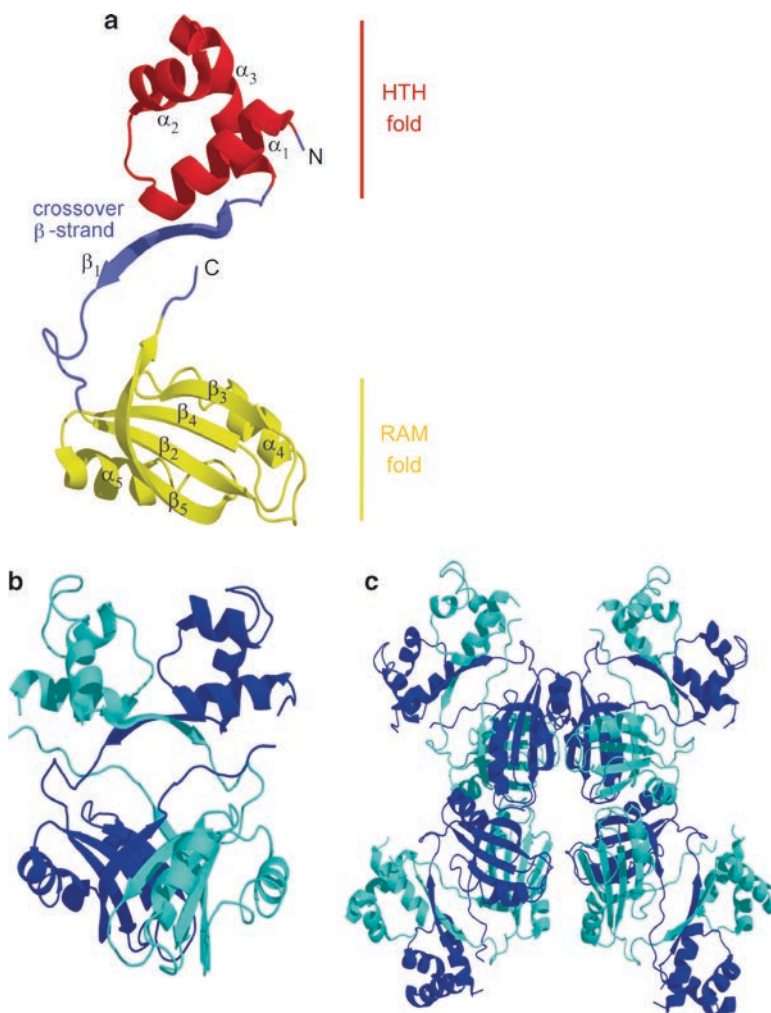


Fig. 8.3 Crystal structure of *E. coli* LRP. (PDB ID: 2GQQ; de los Rios and Perona 2007). The structures of an LRP monomer (a) and dimer (b) are shown. The N-terminal domain contains the HTH motif, coloured in red. The recognition helix (α_3) binds to the major groove of the DNA. The RAM domain (regulation of amino acid metabolism), is coloured in yellow and contains the binding site for the effector leucine. The crossover β -strand involved in dimerization is coloured in blue. The α -helices and β -strands are numbered according to the original publication (de los Rios and Perona 2007). (c) The octameric structure of LRP has been arranged according to the crystallographic data of de los Rios and Perona (2007) by PyMol (DeLano Scientific LLC, South San Francisco, USA)

ousus (Leonard et al. 2001) and LrpC from *Bacillus subtilis* (Thaw et al. 2006). The octameric structure is maintained by hydrophobic interactions of the RAM domains between individual dimer units. In contrast to the octameric structure of *E. coli*

LRP, where the interaction of two of the four LRP dimers is interrupted, the octamers of LrpA and LrpC are closed octameric discs. It is possible, however, that this difference in the quarternary structure is due to the presence of DNA during crystallization of the *E. coli* LRP (de los Rios and Perona 2007).

Binding of LRP to DNA occurs through contacts with the α 3-helix of the HTH motif (see above) and the DNA major groove. Specificity depends on a consensus DNA sequence, which had been determined as (T/C)AG(A/C/T)A(A/T)ATT(A/T)T(A/G/T)CT(G/A) by a SELEX analysis (Cui et al. 1995). For most LRP-regulated operons multiple LRP binding sites have been mapped, to which binding of LRP often occurs with high cooperativity. This is in contrast to the occupation of FIS binding sites, for instance, which are also generally arranged in clusters, but binding of FIS occurs without cooperativity. The cooperative binding mode of LRP often masks the primary binding site and leads to efficient binding of LRP not only at the high-affinity site but includes multiple suboptimal sites. This explains the extended regions of protection within the regulatory regions of many LRP-dependent operons visible in footprint analyses (Wang and Calvo 1993; Nou et al. 1995; Pul et al. 2007). Recent analyses have demonstrated that LRP can bind to non-specific DNA targets with high affinity and in cooperative manner (Peterson et al. 2007). Such cooperative extension after initial binding to a high affinity site appears to be a common phenomenon for LRP-dependent transcription regulation and has specifically been demonstrated in a phasing study, where a synthetic AT-rich sequence had been fused at varying distance upstream of two differentially regulated promoters (Pul et al. 2008). Binding of LRP clearly has an impact on DNA structure. The orientation of the helix-turn-helix-domains on the outside of the octameric LRP core suggests a model in which the DNA is wrapped around the protein core (Thaw et al. 2006). This is consistent with footprint analyses, which for many operons show that LRP-DNA interactions extend over a range of more than 100 base pairs. Moreover, the wrapping of DNA around LrpC has been visualized by electron microscopy and AFM (Beloin et al. 2003). In addition, it has been found that LRP binding results in a periodic pattern of DNase I or hydroxyl radical protected and hypersensitive sites (phased hypersensitivity), indicating LRP-dependent bending or wrapping of the DNA (Stauffer and Stauffer 1994; Ferrario et al. 1995; Nou et al. 1995; Pul et al. 2005; McFarland and Dorman 2008). For LrpC from *B. subtilis* and *E. coli* LRP it has been shown that the formation of such higher-order nucleoprotein complexes constrains supercoils (Beloin et al. 2003; Pul et al. 2007). Interestingly, the reported direction of the constrained supercoils differs for LRP and LrpC. Whether this difference is due to the different quarternary structures of the two homolog proteins or reflects general differences in their function needs further experimental clarification.

It should be noted at this point, that not all members of LRP family are global regulators. *H. influenzae* LrfB shows 75% sequence identity with *E. coli* LRP but has a specific regulatory role restricted to only a few genes (Friedberg et al. 2001). This may be related to the lower expression level of LrfB of only 130 molecules per cell. The authors conclude that the global regulatory function of LRP is restricted to enteric bacteria, which encounter variable environmental conditions. It is likely,

therefore, that in each individual species the function of LRP as a global or a specific regulator is tuned to its expression level. In *E. coli* expression of the *lrp* gene is inversely proportional to the growth rate and coupled to the quality of the media. Moreover, it is induced directly or indirectly through positive regulation by ppGpp (Landgraf et al. 1996). Transcription of the *lrp* gene is enhanced by relaxation of the DNA template following inhibition of DNA gyrase (Müller et al. 2009). Mechanisms to regulate the cellular LRP concentration may thus determine whether it acts as a global or a specific transcription factor.

8.5 IHF: Integration Host Factor

IHF was first described as the host factor required for integration of phage λ in the bacterial chromosome. Today it is known that IHF, in addition to its role in recombination or replication events, is involved in transcription regulation of more than 100 genes in *E. coli* and *S. typhimurium* (Arfin et al. 2000). The heterodimeric IHF in *E. coli* is encoded by the highly homologous genes *ihfA* and *ihfB* (Weisberg et al. 1996), which are translated to the approximately 10 kDa α - and β -subunits, respectively.

IHF binds specifically to the minor groove of DNA by recognizing the consensus sequence (A/T)ATCAANNNTT(G/A) (Hales et al. 1994). This interaction leads to the most pronounced characteristic of IHF, namely its ability to introduce a sharp bend, almost a U-turn, into the DNA. This property makes IHF an important architectural component for many DNA transaction reactions, which require bent DNA conformations. As an example, the structure of *E. coli* IHF bound to the H'-site of phage λ is shown in Fig. 8.4 (Rice et al. 1996). Each of the IHF protein monomers is composed of three α -helices, forming the central body, and two

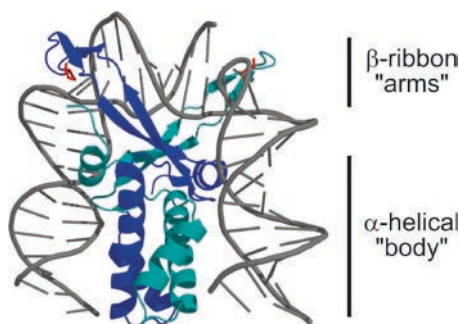


Fig. 8.4 Crystal structure of a DNA-IHF complex (PDB ID: 1IHF; (Rice et al. 1996)). The individual monomers (IHF α and IHF β) are indicated by a blue or green backbone, respectively. The two β -ribbons wrap around the DNA through the minor grooves. The proline residues (shown in red) interrupt the base stacking, leading to strong kinks in the DNA structure, which is shown in grey

antiparallel β -ribbons, forming the IHF-arms. Each of the β -ribbon arms contains a conserved proline residue close to the tip of the antiparallel strands, which, in the bound complex, intercalates between adjacent DNA bases. The intercalation of the two prolines abolishes the stacking interaction between the neighbouring bases, which leads to a widening of the minor groove giving rise to two sharp kinks of the target DNA. Stabilized by electrostatic interactions between positively charged residues of the IHF body and the negatively charged DNA backbone the contour of the DNA, deformed by the two kinks, results in a U-turn like bending of about 160° – 180° (Rice et al. 1996).

A further remarkable aspect of IHF-DNA interaction resides in the sequence-specificity of IHF, which is rather uncommon for minor groove binding proteins. IHF-dependent recognition of the DNA consensus-sequence involves contributions of both side chain-base contacts within the β -arms and also within the IHF body. Thus, the interaction between IHF and DNA is a typical example for the ‘indirect readout’ mechanism for which the recognition and binding to the consensus sequence depends on the structural flexibility of the target DNA (Aeling et al. 2006).

8.6 HU: Histone-Like Protein from Strain U93

The dimeric HU protein exhibits sequence and structural homology with the IHF protein. The first high-resolution structure derived from *B. stearotheophilus* depicts a dimeric molecule with two flexible basic arms, which fit into the minor groove of DNA (White et al. 1999). In most bacteria HU exists as a homodimer. In the enteric bacterium *E. coli*, however, it exists as heterodimer, with each subunit being 9.5 kDa in size, encoded by the homologous *hupA* and *hupB* genes. The subunits, HU α and HU β , show 70% homology to each other and approximately 35% homology to IHF. Despite the sequence and structural similarity between IHF and HU, the latter protein binds sequence non-specific to the DNA but also bends the target DNA through intercalation of two likewise conserved proline residues between two base-pairs in the minor groove (Rice 1997; Bewley et al. 1998; Swinger et al. 2003) (Fig. 8.5). As in the case of IHF the intercalation of proline residues introduces and stabilizes two kinks in the DNA. In a recent binding study it was shown by FRET analysis that interaction of HU with a 34 bp DNA fragment caused almost a 143° bending angle of the DNA (Koh et al. 2008). This result is fully consistent with the resolved structure from co-crystals of *Anabaena* HU and DNA shown in Fig. 8.5.

Consistent with its property to bend DNA HU plays an important architectural role in all kinds of DNA transactions including transcription regulation. In combination with other regulators HU often facilitates or even enables the formation of active DNA conformations through formation of higher-order nucleoprotein complexes. Such DNA structures often involve DNA loops as has been shown for instance for the *gal* operon. In this case HU binds within the interoperator region

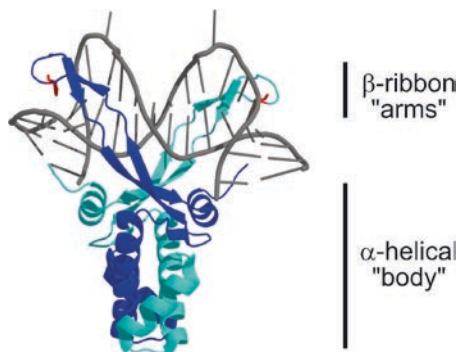


Fig. 8.5 Crystal structure of the *Anabaena* HU-DNA complex (PDB ID: 1P71; Swinger et al. 2003). The individual monomers (HU α and HU β) are shown as blue or green backbones. The structure resembles very much that for IHF. As shown for IHF two conserved proline residues at the tip of the β -ribbons (coloured in red) intercalate between base pairs of the DNA minor groove, inducing a strong DNA bend

of the *gal* operon facilitating the interaction of two bound GalR dimers, which lead to a looped structure inadequate for transcription and therefore called repressosome (Semsey et al. 2002). A similar role for HU has also been described for the site-specific DNA inversion by the Hin recombinase. Here, the Hin-dependent assembly of the invertasome is facilitated by HU, which enables the necessary DNA looping (Haykinson and Johnson 1993) (Chapter 17).

Taking advantage of atomic force microscopy and magnetic tweezers technology some crucial structural properties of HU-DNA complexes have been addressed. Such studies indicated that HU, aside from inducing flexible bends in DNA, is able to form rigid nucleoprotein filaments at higher HU concentrations, indicating the existence of two different HU-DNA nucleoprotein complexes (Dame and Goosen 2002; van Noort et al. 2004). In such filament-like complexes HU is arranged helically around the DNA (see also Fig. 8.7a).

Interestingly, the expression of the two homologous HU subunits is differentially regulated in *E. coli*, leading to diverse subunit compositions of HU in the cell by switching from predominantly HU α homodimers in early log phase to heterodimers at stationary phase (Claret and Rouvière-Yaniv 1997). This phase-dependent expression pattern suggests that the different forms of HU are not functionally equivalent. Rather the distinct dimeric forms exhibit a potential regulatory role in response to different growth conditions (Claret and Rouvière-Yaniv 1997). The ability to form distinct homo- or heterodimers among the NAP members is not restricted to HU. A similar situation has been described for H-NS and the paralog StpA. Both proteins, which are able to interact with each other, are also differentially regulated and disparate functions have been suggested for the distinct dimers (Zhang et al. 1996). The implications of such homomeric or heteromeric complexes between different NAP members for bacterial adaptation and pathogenesis have been discussed in a previous review (Dorman et al. 1999).

8.7 Dps: DNA Binding Protein from Starved Cells

Among the NAP family members Dps is somewhat unique as it does not appear to directly regulate gene expression and seems to exclusively function in DNA compaction and protection against damage and oxidative stress upon induction. The monomeric subunit of the starvation-dependent DNA binding protein Dps has a molecular weight of 19 kDa in *E. coli*. It is the most abundant NAP during the stationary phase with 180,000 molecules per cell (Almirón et al. 1992).

The active form of Dps is a dodecamer (Fig. 8.6b). Together with DNA several such Dps oligomers form hexagonally packed two-dimensional arrays (Almirón

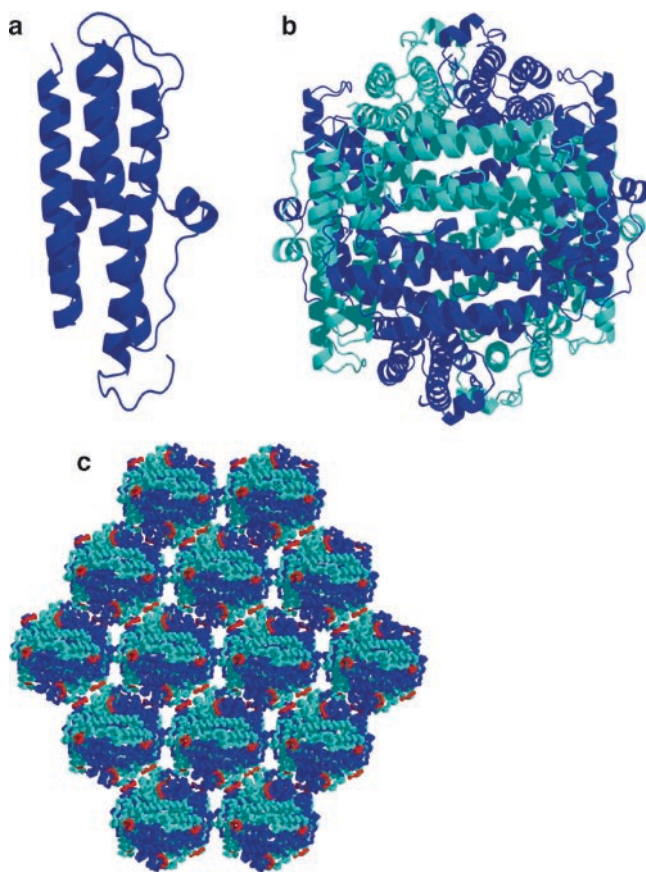


Fig. 8.6 Crystal structure of Dps (PDB ID: 1DPS; Grant et al. 1998). (a) The monomeric Dps structure, largely consisting of a four helix bundle, strongly resembles the structure of ferritin. (b) 12 molecules of Dps are shown in their dodecamer arrangement. (c) A lattice formed by 14 Dps dodecamers, arranged by PyMol is shown. The N-terminal amino acids, which are part of the unstructured lysine-rich domains of all individual Dps monomers are coloured red

et al. 1992). The formation of this tightly packed highly stable crystalline structure protects the DNA in starved *E. coli* cells during stationary phase. The tight packing strongly restricts the accessibility of DNA to damaging agents and protects the nucleoid against oxidative damage, nuclease cleavage, UV light, thermal shock and acidic stress (Almirón et al. 1992; Martinez and Kolter 1997; Choi et al. 2000; Frenkiel-Krispin et al. 2004; Nair and Finkel 2004).

The Dps monomer is a structural homolog of ferritin, a family of iron storage proteins (Grant et al. 1998), consisting of a four helix bundle core (Fig. 8.6a). This homology explains some of the protective functions of Dps against reactive oxygen species (ROS) as outlined below. The monomers associate into a Dps dodecamer with a diameter of approximately 90 Å and a hollow core of ~45 Å in diameter. Dps contains no classical DNA binding motif and the actual DNA binding mechanism of Dps still remains to be elucidated. It is clear, however, that Dps makes non-specific contacts with the DNA. Because the surface of the Dps dodecamer is negatively charged, a simple electrostatic interaction with likewise negatively charged DNA must be excluded (Grant et al. 1998). Rather the lysine-containing disordered N-terminus has been proposed to play an important role for DNA binding. According to this, three dodecamers within the hexagonal Dps crystal lattice form holes, in which the lysine-residues of the N-termini are arranged in lines. It is assumed that the DNA passes through those holes and that the bound DNA becomes stabilized by the basic residues of the disordered N-termini (Grant et al. 1998; Ceci et al. 2004). Actually, Dps variants without positively charged N-termini are impeded in DNA binding. Moreover, studies with *E. coli* Dps deletion mutants lacking the lysine-rich N-terminus confirmed the essential role of Dps in self-aggregation and DNA compaction (Ceci et al. 2004). Figure 8.6c shows a model of a lattice of Dps dodecamer in which the disordered N-termini of each Dps monomeric subunits are indicated in red.

Dps not only protects from DNA damage by formation of a crystalline structure with DNA, thereby dramatically reducing the accessibility of vulnerable DNA positions. It also actively protects against reactive oxygen radicals. In accordance with its structural similarity to ferritin, Dps is able to inhibit Fe^{2+} -dependent generation of free radicals by the Fenton reaction through binding, sequestering and oxidation of Fe^{2+} . Protection can take place, whether or not Dps is bound to DNA (Ceci et al. 2003, 2004). These studies suggest that Dps in its non-DNA-bound form is still involved in the defence of bacterial pathogens by protection against H_2O_2 , which is produced by the host defence system. Binding of Dps to DNA seems to be responsible for nucleoid condensation and protection against other damaging agents or factors, such as low pH, nucleases or UV radiation. In line with the notion of being involved in acid resistance it has been shown that *E. coli* Dps binds with higher affinity to DNA at low pH conditions (Ceci et al. 2004). In summary, Dps plays an important role in protecting bacteria against different types of long-time stress. Its major function consists in avoiding DNA damage under the harmful conditions, which the cells encounter during starvation.

In contrast to the other nucleoid-associated proteins no specific function of Dps as transcription factor with regulatory properties has as yet been described. There

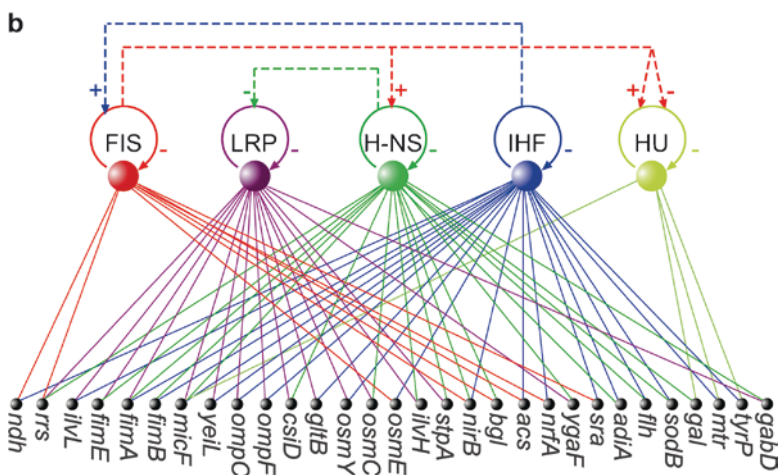
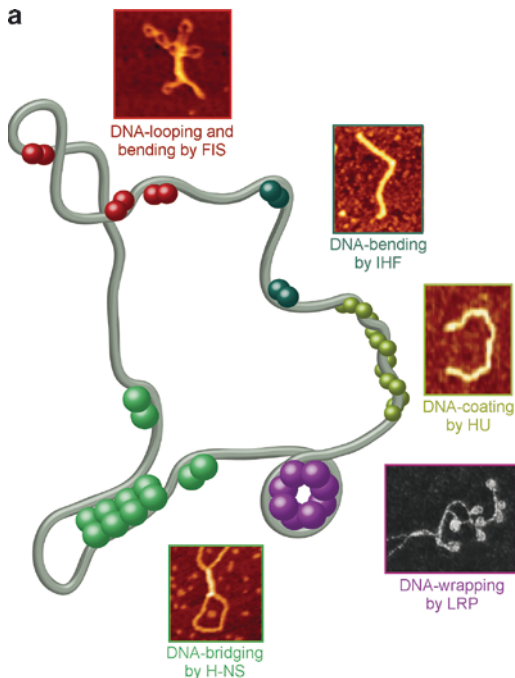


Fig. 8.7 DNA-binding and regulatory properties of NAPs. **(a)** Different modes of DNA structuring by NAPs are schematically illustrated. For each NAP a schematic drawing and a corresponding AFM image is shown. Different effects of NAP binding on the target DNA are displayed: plectonemic loop (FIS), sharp DNA bends (IHF), DNA-coating (HU), DNA-wrapping (LRP) and DNA-bending and bridging (H-NS) (Taken with permission from Pul and Wagner (2007)). The original scanning force microscopy images of FIS and H-NS are courtesy of G. Muskhelishvili and R.T. Dame, respectively. The IHF, H-NS and LRP pictures were originally published by Luijsterburg et al. (2006), the HU picture by van Noort et al. (2004). **(b)** The regulatory network of genes known to be under the control of the NAPs FIS, LRP, H-NS, IHF and HU, is presented. Only those genes are indicated which are controlled by at least two NAP members. The data have been taken from Regulon DB (2008). In addition, coloured arrows indicate autoregulation and dashed arrows cross-regulation of the NAPs. In case of HU, FIS activates the expression of HU α (*hupA*) and represses transcription of HU β (*hupB*) (Claret and Rouvière-Yaniv 1996)

are indications from recent studies, however, that Dps may affect replication through interaction with the DnaA protein, interfering with DnaA-dependent unwinding of *oriC* (Chodavarapu et al. 2008).

8.8 Conclusions

All the members of the family of NAPs, especially those exemplified in this review, share the unique properties of compacting DNA and dynamically modulating the structure of bacterial nucleoids. Many of the NAPs affect DNA supercoiling, which contributes to DNA compaction, but also is essential for different kinds of DNA transactions, such as replication, recombination or transcription. As such, NAPs play an important role in the maintenance of a dynamic genome and directly or indirectly affect gene expression. Their expression levels or activities in the cell often depend on environmental stimuli or changing growth conditions, which is consistent with their task as environmental sensors. Their mode of interaction with DNA is quite variable. The specificity ranges from highly sequence-specific to completely non-specific. Binding often involves oligomeric forms of the NAPs and interaction of different NAPs with DNA generally results in defined three-dimensional structures. According to their effects on DNA structure NAPs can be generally categorized in ‘benders’, ‘bridgers’ or ‘wrappers’ (Luijsterburg et al. 2006, 2008). This is illustrated in Fig. 8.7a. Another characteristic property of all NAPs consists in their capability to act in concert and to form distinct hetero- or homomeric protein complexes. Often NAPs share overlapping DNA binding sites, indicative of synergistic or antagonistic functions. They affect expression of each other and together coordinate the regulation of a large number of different genes (Fig. 8.7b).

Together, the importance of NAPs as global regulators of the bacterial cell has become more and more evident. They are involved in transcription regulation through many different mechanisms and often in cooperation with each other or with different gene-specific regulatory factors. This combinatorial regulation leads to integration of different external or internal signals and contributes to the efficient adaptation of the cell to environmental changes.

References

- Aeling KA, Opel ML, Steffen NR, Tretyachenko-Ladokhina V, Hatfield GW, Lathrop RH, Senear DF (2006) Indirect recognition in sequence-specific DNA binding by *Escherichia coli* integration host factor: the role of DNA deformation energy. *J Biol Chem* 281:39236–39248
- Afflerbach H, Schröder O, Wagner R (1999) Conformational changes of the upstream DNA mediated by H-NS and FIS regulate *E. coli rrmB* P1 promoter activity. *J Mol Biol* 286:339–353
- Almirón M, Link AJ, Furlong D, Kolter R (1992) A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev* 6:2646–2654

- Arfin SM, Long AD, Ito ET, Toller L, Riehle MM, Paegle ES, Hatfield GW (2000) Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor. *J Biol Chem* 275:29672–29684
- Badaut C, Williams R, Arlusion V, Bouffartigues E, Robert B, Buc H, Rimsky S (2002) The degree of oligomerization of the H-NS nucleoid structuring protein is related to specific binding to DNA. *J Biol Chem* 277:41657–41666
- Beloin C, Jeusset J, Revet B, Mirambeau G, Le Hegarat F, Le Cam E (2003) Contribution of DNA conformation and topology in right-handed DNA wrapping by the *Bacillus subtilis* LrpC protein. *J Biol Chem* 278:5333–5342
- Bertin P, Benhabiles N, Krin E, Laurent-Winter C, Tendeng C, Turlin E, Thomas A, Danchin A, Brasseur R (1999) The structural and functional organization of H-NS-like proteins is evolutionarily conserved in Gram-negative bacteria. *Mol Microbiol* 31:319–329
- Bewley CA, Gronenborn AM, Clore GM (1998) Minor groove-binding architectural proteins: structure, function, and DNA recognition. *Annu Rev Biophys Biomol Struct* 27:105–131
- Bloch V, Yang Y, Margeat E, Chavanieu A, Auge MT, Robert B, Arold S, Rimsky S, Kochoyan M (2003) The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat Struct Biol* 10:212–218
- Bokal AJ, Ross W, Gaal T, Johnson RC, Gourse RL (1997) Molecular anatomy of a transcription activation patch: FIS-RNA polymerase interactions at the *Escherichia coli* *rrnB* P1 promoter. *EMBO J* 16:154–162
- Bouffartigues E, Buckle M, Badaut C, Travers A, Rimsky S (2007) H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat Struct Mol Biol* 14:441–448
- Brandi A, Spurio R, Gualerzi CO, Pon CL (1999) Massive presence of the *Escherichia coli* ‘major cold-shock protein’ CspA under non-stress conditions. *EMBO J* 18:1653–1659
- Calvo JM, Matthews RG (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* 58:466–490
- Ceci P, Ilari A, Falvo E, Chiancone E (2003) The Dps protein of *Agrobacterium tumefaciens* does not bind to DNA but protects it toward oxidative cleavage: x-ray crystal structure, iron binding, and hydroxyl-radical scavenging properties. *J Biol Chem* 278:20319–20326
- Ceci P, Cellai S, Falvo E, Rivetti C, Rossi GL, Chiancone E (2004) DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Res* 32:5935–5944
- Ceschini S, Lupidi G, Coletta M, Pon CL, Fioretti E, Angeletti M (2000) Multimeric self-assembly equilibria involving the histone-like protein H-NS. A thermodynamic study. *J Biol Chem* 275:729–734
- Chen S, Calvo JM (2002) Leucine-induced dissociation of *Escherichia coli* Lrp hexadecamers to octamers. *J Mol Biol* 318:1031–1042
- Chen S, Rosner MH, Calvo JM (2001) Leucine-regulated self-association of leucine-responsive regulatory protein (Lrp) from *Escherichia coli*. *J Mol Biol* 312:625–635
- Chen S, Iannolo M, Calvo JM (2005) Cooperative binding of the leucine-responsive regulatory protein (Lrp) to DNA. *J Mol Biol* 345:251–264
- Cheng YS, Yang WZ, Johnson RC, Yuan HS (2000) Structural analysis of the transcriptional activation region on Fis: crystal structures of six Fis mutants with different activation properties. *J Mol Biol* 302:1139–1151
- Cho BK, Knight EM, Barrett CL, Palsson BO (2008) Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Res* 18:900–910
- Chodavarapu S, Gomez R, Vicente M, Kaguni JM (2008) *Escherichia coli* Dps interacts with DnaA protein to impede initiation: a model of adaptive mutation. *Mol Microbiol* 67:1331–1346
- Choi SH, Baumler DJ, Kaspar CW (2000) Contribution of *dps* to acid stress tolerance and oxidative stress tolerance in *Escherichia coli* O157:H7. *Appl Environ Microbiol* 66:3911–3916
- Claret L, Rouvière-Yaniv J (1997) Variation in HU composition during growth of *Escherichia coli*: heterodimer is required for long term survival. *J Mol Biol* 273:93–104
- Claret L, Rouvière-Yaniv J (1996) Regulation of HU α and HU β by CRP and FIS in *Escherichia coli*. *J Mol Biol* 263:126–139

- Colland F, Barth M, Hengge-Aronis R, Kolb A (2000) Sigma factor selectivity of *Escherichia coli* RNA polymerase: role for CRP, IHF and Lrp transcription factors. *EMBO J* 19:3028–3037
- Cui Y, Wang Q, Stormo GD, Calvo JM (1995) A consensus sequence for binding of Lrp to DNA. *J Bacteriol* 177:4872–4880
- Dame RT, Goosen N (2002) HU: promoting or counteracting DNA compaction? *FEBS Lett* 529:151
- Dame RT, Wyman C, Goosen N (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* 28:3504–3510
- Dame RT, Wyman C, Wurm R, Wagner R, Goosen N (2002) Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at *rnnB* P1. *J Biol Chem* 277:2146–2150
- Dame RT, Luijsterburg MS, Krin E, Bertin PN, Wagner R, Wuite GJ (2005) DNA Bridging: a property shared among H-NS-like proteins. *J Bacteriol* 187:1845–1848
- Dame RT, Noom MC, Wuite GJ (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* 444:387–390
- de los Rios S, Perona JJ (2007) Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *J Mol Biol* 366:1589–1602
- Donato GM, Kawula TH (1998) Enhanced binding of altered H-NS protein to flagellar rotor protein FliG causes increased flagellar rotational speed and hypermotility in *Escherichia coli*. *J Biol Chem* 273:24030–24036
- Dorman CJ (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* 2:391–400
- Dorman CJ (2007) H-NS, the genome sentinel. *Nat Rev Microbiol* 5:157–161
- Dorman CJ, Hinton JC, Free A (1999) Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends Microbiol* 7:124–128
- Esposito D, Gerard GF (2003) The *Escherichia coli* Fis protein stimulates bacteriophage lambda integrative recombination in vitro. *J Bacteriol* 185:3076–3080
- Esposito D, Petrovic A, Harris R, Ono S, Eccleston JF, Mbabaali A, Haq I, Higgins CF, Hinton JC, Driscoll PC, Ladbury JE (2002) H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *J Mol Biol* 324:841–850
- Ettema TJ, Brinkman AB, Tani TH, Rafferty JB, van der Oost J (2002) A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes. *J Biol Chem* 277:37464–37468
- Falconi M, Brandi A, La Teana A, Gualerzi CO, Pon CL (1996) Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of *hns* expression. *Mol Microbiol* 19:965–975
- Falconi M, Colonna B, Prosseda G, Micheli G, Gualerzi CO (1998) Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. *EMBO J* 17:7033–7043
- Fang FC, Rimsky S (2008) New insights into transcriptional regulation by H-NS. *Curr Opin Microbiol* 11:113–120
- Feldman-Cohen LS, Shao Y, Meinhold D, Miller C, Colon W, Osuna R (2006) Common and variable contributions of Fis residues to high-affinity binding at different DNA sequences. *J Bacteriol* 188:2081–2095
- Ferrario M, Ernsting BR, Borst DW, Wiese DE, Blumenthal RM, Matthews RG (1995) The leucine-responsive regulatory protein of *Escherichia coli* negatively regulates transcription of *ompC* and *micF* and positively regulates translation of *ompF*. *J Bacteriol* 177:103–113
- Frenkiel-Krispin D, Ben-Avraham I, Englander J, Shimoni E, Wolf SG, Minsky A (2004) Nucleoid restructuring in stationary-state bacteria. *Mol Microbiol* 51:395–405
- Friedberg D, Midkiff M, Calvo JM (2001) Global versus local regulatory roles for Lrp-related proteins: *Haemophilus influenzae* as a case study. *J Bacteriol* 183:4004–4011
- Gama-Castro S, Jimenez-Jacinto V, Peralta-Gil M, Santos-Zavaleta A, Penalzoza-Spinola MI, Contreras-Moreira B, Segura-Salazar J, Muniz-Rascado L, Martinez-Flores I, Salgado H, Bonavides-Martinez C, Abreu-Goodger C, Rodriguez-Penagos C, Miranda-Rios J, Morett E, Merino E, Huerta AM, Trevino-Quintanilla L, Collado-Vides J (2008) RegulonDB (version 6.0): gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nucleic Acids Res* 36(Database issue): D120–4

- Grainger DC, Hurd D, Goldberg MD, Busby SJ (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res* 34:4642–4652
- Grainger DC, Goldberg MD, Lee DJ, Busby SJ (2008) Selective repression by Fis and H-NS at the *Escherichia coli* *dps* promoter. *Mol Microbiol* 68:1366–1377
- Grant RA, Filman DJ, Finkel SE, Kolter R, Hogle JM (1998) The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat Struct Biol* 5:294–303
- Hales LM, Gumpert RI, Gardner JF (1994) Determining the DNA sequence elements required for binding integration host factor to two different target sites. *J Bacteriol* 176:2999–3006
- Haykinson MJ, Johnson RC (1993) DNA looping and the helical repeat in vitro and in vivo: effect of HU protein and enhancer location on Hin invertasome assembly. *EMBO J* 12:2503–2512
- Hillebrand A, Wurm R, Menzel A, Wagner R (2005) The seven *E. coli* ribosomal RNA operon upstream regulatory regions differ in structure and transcription factor binding efficiencies. *Biol Chem* 386:523–534
- Hirvonen CA, Ross W, Wozniak CE, Marasco E, Anthony JR, Aiyar SE, Newburn VH, Gourse RL (2001) Contributions of UP elements and the transcription factor FIS to expression from the seven *rrm* P1 promoters in *Escherichia coli*. *J Bacteriol* 183:6305–6314
- Kajitani M, Ishihama A (1991) Identification and sequence determination of the host factor gene for bacteriophage Q beta. *Nucleic Acids Res* 19:1063–1066
- Koh J, Saecker RM, Record MT Jr (2008) DNA binding mode transitions of *Escherichia coli* HU (alphabet): evidence for formation of a bent DNA – protein complex on intact, linear duplex DNA. *J Mol Biol* 383:324–346
- Kostrewa D, Granzin J, Koch C, Choe HW (1991) Three-dimensional structure of the *E. coli* DNA-binding protein Fis. *Nature* 349:178–180
- Kostrewa D, Granzin J, Stock D, Choe H-W, Labahn J, Saenger W (1992) Crystal structure of the factor for inversion stimulation FIS at 2 Å resolution. *J Mol Biol* 226:209–226
- Landgraf JR, Wu J, Calvo JM (1996) Effects of nutrition and growth rate on Lrp levels in *Escherichia coli*. *J Bacteriol* 178:6930–6936
- Lang B, Blot N, Bouffartigues E, Buckle M, Geertz M, Gualerzi CO, Mavathur R, Muskhelishvili G, Pon CL, Rimsky S, Stella S, Babu MM, Travers A (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* 35:6330–6337
- Leonard PM, Smits SH, Sedelnikova SE, Brinkman AB, de Vos WM, van der Oost J, Rice DW, Rafferty JB (2001) Crystal structure of the Lrp-like transcriptional regulator from the archaeon *Pyrococcus furiosus*. *EMBO J* 20:990–997
- Liu Q, Richardson CS (1993) Gene 5.5 protein of bacteriophage T7 inhibits the nucleoid protein H-NS of *Escherichia coli*. *Proc Natl Acad Sci USA* 90:1761–1765
- Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M, Hinton JC (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog* 2:e81
- Luijsterburg MS, Noom MC, Wuite GJ, Dame RT (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol* 156:262–272
- Luijsterburg MS, White MF, van Driel R, Dame RT (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol* 43:393–418
- Martinez A, Kolter R (1997) Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J Bacteriol* 179:5188–5194
- McFarland KA, Dorman CJ (2008) Autoregulated expression of the gene coding for the leucine-responsive protein, Lrp, a global regulator in *Salmonella enterica* serovar Typhimurium. *Microbiology* 154:2008–2016
- Müller CM, Aberg A, Straseviciene J, Emody L, Uhlin BE, Balsalobre C (2009) Type 1 fimbriae, a colonization factor of uropathogenic *Escherichia coli*, are controlled by the metabolic sensor CRP-cAMP. *PLoS Pathog* 5:e1000303

- Muskhelishvili G, Buckle M, Heumann H, Kahmann R, Travers AA (1997) FIS activates sequential steps during transcription initiation at a stable RNA promoter. *EMBO J* 12:3655–3665
- Nair S, Finkel SE (2004) Dps protects cells against multiple stresses during stationary phase. *J Bacteriol* 186:4192–4198
- Navarre WW, McClelland M, Libby SJ, Fang FC (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev* 21:1456–1471
- Nieto JM, Madrid C, Miquelay E, Parra JL, Rodriguez S, Juarez A (2002) Evidence for direct protein-protein interaction between members of the enterobacterial Hha/YmoA and H-NS families of proteins. *J Bacteriol* 184:629–635
- Nilsson L, Vanet A, Vijgenboom E, Bosch L (1990) The role of Fis in *trans* activation of stable RNA operons of *E. coli*. *EMBO J* 9:727–734
- Nilsson L, Verbeek H, Vijgenboom E, van Drunen C, Vanet A, Bosch L (1992) Fis-dependent *trans*-activation of stable RNA operons of *Escherichia coli* under various growth conditions. *J Bacteriol* 174:921–929
- Nou X, Braaten B, Kaltenbach L, Low DA (1995) Differential binding of Lrp to two sets of pap DNA binding sites mediated by Pap I regulates Pap phase variation in *Escherichia coli*. *EMBO J* 14:5785–5797
- Pan CQ, Feng JA, Finkel SE, Landgraf R, Sigman D, Johnson RC (1994) Structure of the *Escherichia coli* Fis-DNA complex probed by protein conjugated with 1, 10-phenanthroline copper (I) complex. *Proc Natl Acad Sci USA* 91:1721–1725
- Pan CQ, Finkel SE, Cramton SE, Feng J-A, Sigman DS, Johnson RC (1996) Variable structures of Fis-DNA complexes determined by flanking DNA – protein contacts. *J Mol Biol* 264:675–695
- Peterson SN, Dahlquist FW, Reich NO (2007) The role of high affinity non-specific DNA binding by Lrp in transcriptional regulation and DNA organization. *J Mol Biol* 369:1307–1317
- Platko JV, Calvo JM (1993) Mutations affecting the ability of *Escherichia coli* Lrp to bind DNA, activate transcription, or respond to leucine. *J Bacteriol* 175:1110–1117
- Pul Ü, Wagner R (2007) DNA-binding proteins: nucleoid associated proteins in microorganisms. *BioSpektrum* 13:495–497
- Pul Ü, Wurm R, Lux B, Meltzer M, Menzel A, Wagner R (2005) LRP and H-NS – cooperative partners for transcription regulation at *E. coli* rRNA promoters. *Mol Microbiol* 58:864–876
- Pul Ü, Wurm R, Wagner R (2007) The role of LRP and H-NS in transcription regulation: involvement of synergism, allostery and macromolecular crowding. *J Mol Biol* 366:900–915
- Pul Ü, Lux B, Wurm R, Wagner R (2008) Effect of upstream curvature and transcription factors H-NS and LRP on the efficiency of *Escherichia coli* rRNA promoters P1 and P2 – a phasing analysis. *Microbiology* 154:2546–2558
- Renzoni D, Esposito D, Pfuhl M, Hinton JC, Higgins CF, Driscoll PC, Ladbury JE (2001) Structural characterization of the N-terminal oligomerization domain of the bacterial chromatin-structuring protein, H-NS. *J Mol Biol* 306:1127–1137
- Rice PA (1997) Making DNA do a U-turn: IHF and related proteins. *Curr Opin Struct Biol* 7:86–93
- Rice PA, Yang S-W, Mizuuchi K, Nash H (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* 27:1295–1306
- Rimsky S (2004) Structure of the histone-like protein H-NS and its role in regulation and genome superstructure. *Curr Opin Microbiol* 7:109–114
- Schneider R, Lurz R, Luder G, Tolksdorf C, Travers A, Muskhelishvili G (2001) An architectural role of the *Escherichia coli* chromatin protein Fis in organising DNA. *Nucleic Acids Res* 29:5107–5114
- Schröder O, Wagner R (2000) The bacterial DNA-binding protein H-NS represses ribosomal RNA transcription by trapping RNA polymerase in the initiation complex. *J Mol Biol* 298:737–748
- Schröder O, Wagner R (2002) The bacterial regulatory protein H-NS – a versatile modulator of nucleic acid structures. *Biol Chem* 383:945–960
- Schröder O, Tippner D, Wagner R (2001) Towards the three-dimensional structure of the *E. coli* DNA-binding protein H-NS: a CD- and fluorescence study. *Biochem Biophys Res Commun* 282:219–227

- Semsey S, Geanakopoulou M, Lewis DE, Adhya S (2002) Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *EMBO J* 21:4349–4356
- Sette M, Spurio R, Trotta E, Brandizi C, Brandi A, Pon CL, Barbato G, Boelens R, Gualerzi CO (2009) Sequence-specific recognition of DNA by the C-terminal domain of *Escherichia coli* nucleoid-associated protein H-NS. *J Biol Chem* 284:30453–30462
- Shao Y, Feldman-Cohen LS, Osuna R (2008a) Biochemical identification of base and phosphate contacts between Fis and a high-affinity DNA binding site. *J Mol Biol* 380:327–339
- Shao Y, Feldman-Cohen LS, Osuna R (2008b) Functional characterization of the *Escherichia coli* Fis-DNA binding sequence. *J Mol Biol* 376:771–785
- Shin M, Song M, Rhee JH, Hong Y, Kim YJ, Seok YJ, Ha KS, Jung SH, Choy HE (2005) DNA looping-mediated repression by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping. *Genes Dev* 19:2388–2398
- Shindo H, Iwaki T, Ieda R, Kurumizaka H, Ueguchi C, Mizuno T, Morikawa S, Nakamura H, Kuboniwa H (1995) Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from *Escherichia coli*. *FEBS Lett* 360:125–131
- Shindo H, Ohnuki A, Ginba H, Katoh E, Ueguchi C, Mizuno T, Yamazaki T (1999) Identification of the DNA binding surface of H-NS protein from *Escherichia coli* by heteronuclear NMR spectroscopy. *FEBS Lett* 455:63–69
- Skoko D, Yoo D, Bai H, Schnurr B, Yan J, McLeod SM, Marko JF, Johnson RC (2006) Mechanism of chromosome compaction and looping by the *Escherichia coli* nucleoid protein Fis. *J Mol Biol* 364:777–798
- Spurio R, Falconi M, Brandi A, Pon CL, Gualerzi CO (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *EMBO J* 16:1795–1805
- Stauffer LT, Stauffer GV (1994) Characterization of the *gcv* control region from *Escherichia coli*. *FEMS Microbiol Lett* 131:95–98
- Stella S, Spurio R, Falconi M, Pon CL, Gualerzi CO (2005) Nature and mechanism of the in vivo oligomerization of nucleoid protein H-NS. *EMBO J* 24:2896–2905
- Stoebel DM, Free A, Dorman CJ (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* 154:2533–2545
- Swinger KK, Lemberg KM, Zhang Y, Rice PA (2003) Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J* 22:3749–3760
- Tani TH, Khodursky A, Blumenthal RM, Brown PO, Matthews RG (2002) Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci USA* 99:13471–13476
- Thaw P, Sedelnikova SE, Muranova T, Wiese S, Ayora S, Alonso JC, Brinkman AB, Akerboom J, van der Oost J, Rafferty JB (2006) Structural insight into gene transcriptional regulation and effector binding by the Lrp/AsnC family. *Nucleic Acids Res* 34:1439–1449
- Turner EC, Dorman CJ (2007) H-NS antagonism in *Shigella flexneri* by VirB, a virulence gene transcription regulator that is closely related to plasmid partition factors. *J Bacteriol* 189:3403–3413
- Ueguchi C, Suzuki T, Yoshida T, Tanaka K, Mizuno T (1996) Systematic mutational analysis revealing the functional domain organization of *Escherichia coli* nucleoid protein H-NS. *J Mol Biol* 263:149–162
- van Noort J, Verbrugge S, Goosen N, Dekker C, Dame RT (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc Natl Acad Sci USA* 101:6969–6974
- Varshavsky AJ, Nedospasov SA, Bakajeva VV, Georgiev GP (1977) Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Res* 4:2725–2745
- Wang Q, Calvo JM (1993) Lrp, a global regulatory protein of *Escherichia coli*, binds co-operatively to multiple sites and activates transcription of *ilvIH*. *J Mol Biol* 229:306–318
- White SW, Wilson KS, Appelt K, Tanaka I (1999) The high-resolution structure of DNA-binding protein HU from *Bacillus stearothermophilus*. *Acta Crystallogr D Biol Crystallogr* 55:801–809
- Weisberg RA, Freundlich M, Friedman D, Gardner J, Goosen N, Nash, H Oppenheim A, Rouvière-Yaniv J (1996) Nomenclature of the genes encoding IHF. *Mol Microbiol* 19:642

- Williams RM, Rimsky S, Buc H (1996) Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. *J Bacteriol* 178:4335–4343
- Yamada H, Muramatsu S, Mizuno T (1990) An *Escherichia coli* protein that preferentially binds to sharply curved DNA. *J Biochem (Tokyo)* 108:420–425
- Yokoyama K, Ishijima SA, Clowney L, Koike H, Aramaki H, Tanaka C, Makino K, Suzuki M (2006) Feast/famine regulatory proteins (FFRPs): *Escherichia coli* Lrp, AsnC and related archaeal transcription factors. *FEMS Microbiol Rev* 30:89–108
- Zhang A, Rimsky S, Reaban ME, Buc H, Belfort M (1996) *Escherichia coli* protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *EMBO J* 15:1340–1349

Chapter 9

Dps and Bacterial Chromatin

Hanne Ingmer

Abstract When Dps was first discovered in *Escherichia coli* it was soon realized that it is an unusual protein. The purified Dps bound efficiently, but non-specifically, to DNA forming highly ordered complexes and mutants lacking Dps were sensitive to oxidative stress. Since then Dps proteins have been found in many bacteria. Structural studies revealed that the Dps monomer forms a bundle structure resembling the iron binding ferritins and bacterioferritins and upon oligomerization it assembles into a dodecamer with a hollow core. Like ferritins Dps proteins are also able to sequester and detoxify iron through a ferroxidase center that uniquely to the Dps family is shared between two monomers. Thus, the protective capabilities of Dps rely both on its ability to bind and physically protect DNA and its ability to detoxify iron that otherwise may catalyze the production of toxic free radicals. On the other hand Dps can also be used for iron storage under iron limiting conditions. Iron restriction and oxidative stress characterizes the environment that bacterial pathogens encounter in the human host and Dps proteins are required for full virulence of several pathogens. Thus, Dps is a versatile protein that at multiple levels protects bacterial cells against stress.

Keywords Dps • oxidative stress • crystallization • ferroxidase

9.1 Introduction

Oxidative stress is a universal phenomenon that can be experienced by both aerobic and anaerobic micro-organisms. Particularly exposed are the aerobic bacteria that generate reactive oxygen species (ROS) as part of respiration. ROS may damage cellular macromolecules including proteins, lipids and DNA and bacterial pathogens

H. Ingmer (✉)

Department of Veterinary Disease Biology, Faculty of Life Sciences, University of Copenhagen, Stigbøjlen 4, DK-1870, Frederiksberg C, Denmark
e-mail: hi@life.ku.dk

encounter ROS when they come into contact with the human host (Farr and Kogoma 1991; Storz et al. 1990; Miller and Britigan 1997). While the toxicity of some reactive oxygen molecules, like H_2O_2 , is relatively weak they can be transformed into highly active hydroxyl radicals ($\text{OH}\cdot$) in the presence of transition metals such as Fe^{2+} (Fe(II)) according to the Fenton reaction: $\text{H}_2\text{O}_2 + \text{Fe}^{2+} \rightarrow \text{OH}\cdot + \text{OH}^- + \text{Fe}^{3+}$ (Henle and Linn 1997; Luo et al. 1994). Thus, the severity of oxidative stress is highly dependent on the availability of free iron. On the other hand iron is an essential co-factor in many proteins participating in a variety of cellular processes such as respiration, the TCA cycle, oxygen transport and DNA biosynthesis (Crosa 1997). In order to cope with this dilemma, eukaryotic and prokaryotic cells have adopted proteins capable of sequestering, storing and detoxifying free extracellular iron such as ferritins, transferrin and lactoferrin (Hartford et al. 1993; Wooldridge and Williams 1993). In bacterial cells two major families of proteins were originally recognized as being involved in sequestering iron, namely the heme-containing bacterioferritins and the non-heme ferritins (Andrews 1998). However since the early 1990s, a third class of proteins has been recognized as possessing a ferritin-like function providing both iron and hydrogen peroxide detoxification properties. This class of proteins is termed Dps for DNA protection during starvation. Subsequently these proteins have attracted particular interest as they structurally resemble the ferritins but provide iron detoxification through a uniquely located ferroxidase center shared by two Dps monomers and because many members have been shown to bind directly to DNA and physically compact and protect the DNA from oxidative damage. In this chapter the focus will be on the biological role of eubacterial Dps proteins in bacterial stress protection and from the conditions that control the expression of Dps the reader will get some insight into the environmental niches where Dps is important.

9.2 Dps Structure

Early in the study of the *Escherichia coli* Dps protein it was realized that an unusual protein had been discovered. The purified protein bound efficiently to DNA and formed stable complexes that resisted heat treatment and was visualized by electron microscopy as large, highly ordered two-dimensional arrays described as honeycombs (Almirón et al. 1992). Six years later the crystal structure was solved at 1.6 Å resolution by X-ray crystallography revealing that the *E. coli* Dps monomer displays the same four-helix bundle structure as ferritin and bacterioferritin but upon oligomerization it assembles into a ball-like dodecamer with a hollow core and pores at three-fold axes of symmetry in contrast to the 24-mer of ferritins (Grant et al. 1998). In other bacterial species the spherical shape of the dodecamer has consistently been reproduced for a number of Dps family member proteins (Tonello et al. 1999; Ilari et al. 2000; Zanotti et al. 2002; Stillman et al. 2005; Kim et al. 2006).

Little is known about the oligomerization process of the Dps monomers into the dodecamer. For one of the Dps paralogues present in *Mycobacterium smegmatis*, MsDps1, a C-terminal extension of 26 C-terminal amino acids is required for oligomerization (Roy et al. 2007) whereas in the other paralogue, MsDps2,

the N-terminal extension stabilizes the dodecamer (Roy et al. 2008). However, the process of oligomerization appears not to be the same in all organisms. In *E. coli* Dps cross-linking with glutaraldehyde predominantly forms trimers and dodecamers (Grant et al. 1998) and similarly the *Mycobacterium smegmatis* Dps shifts between a trimeric and dodecameric form depending on environmental conditions (Gupta and Chatterji 2003). When the same type of experiment was performed with *Deinococcus radiodurans* Dps it formed dimers and dodecamers (Grove and Wilkinson 2005). The dodecamer assembly for the *D. radiodurans* Dps-1 requires an N-terminal metal binding site (Bhattacharyya and Grove 2007) that has to date only been identified in two other Dps paralogues namely those of *Lactococcus lactis* (Stillman et al. 2005). The importance of the iron binding for stabilization of the dodecamer was recently further emphasized when the amino acids involved in iron binding of the *Helicobacter pylori* Dps protein, HP-NAP were substituted with alanine. The resulting mutant proteins were unable to initiate the formation of stable dimers that for this protein are a prerequisite for dodecamer formation (Kottakis et al. 2008).

Along with the structural studies it was also realized that the traditional ferroxidase centre normally identified within the four-helix bundle of the individual ferritin subunit is not present in Dps proteins confirming that the Dps proteins form a family truly distinct from the ferritins (Grant et al. 1998; Ilari et al. 2000). Rather, in Dps the di-nuclear ferroxidase centers are located at the interfaces of the Dps subunits with the iron binding being coordinated between highly conserved Glu, Asp and His residues as shown in Fig. 9.1 (Ilari et al. 2000, 2002; Zanotti et al. 2002; Zhao et al. 2002; Ren et al. 2003; Kauko et al. 2006).

Members of the Dps family of protein are found in many eubacteria as well as in some archaeobacteria (Wiedenheft et al. 2005; Ramsay et al. 2006; Morikawa et al. 2006; Ohniwa et al. 2006) and in many cases they reside alongside members of the ferritins and the bacterioferritins. Generally, the primary sequence of Dps does not show homology to other *E. coli* nucleoid-associated proteins such as HU or H-NS. However Grant et al. (1998) noted some resemblance to pilin protein. Also the Dps proteins do not display homology to the ferritins but they show limited homology to the bacterioferritins, which has led to the suggestion that they are divergent members of the bacterioferritin/ferritin superfamily (Peña and Bullerjahn 1995; Evans et al. 1995; Bozzi et al. 1997). In a recent bioinformatical analysis it was found that a total of 245 genome-sequenced bacteria contained Dps and 50 of these contained two or more homologues (Roy et al. 2008). The signature sequence defining the Dps family was described as L(X)₁₇-HW(X)₃-G(X)₆-H(X)₁₄-D(X)₃-ER(X)₅₉₋₆₁-D(X)₁₈-(W/H) and in most family members these residues are conserved with notable exceptions such as the *L. lactis* DpsB (Fig. 9.1). Furthermore it was realized that important pathogens like *Mycobacterium leprae*, *M. tuberculosis*, *Neisseria meningitis* and *Chlamydia pneumoniae* do not possess any Dps-like proteins (Roy et al. 2008). Interestingly, in the lactic acid bacteria the only ferritin-like proteins expressed are members of the Dps family (Yamamoto et al. 2002). These organisms rely solely on fermentation for energy production and are characterized by being unable to synthesize heme or produce catalase. Consequently many lactic acid bacteria are sensitive to oxygen and oxidative stress but despite these deficiencies they have only been equipped with one type of ferritin-like protein, highlighting the potential of Dps in oxidative stress protection.

E-coli	MSTAKLVKSKATNLLYTRNDVSDSEKKATVELLNRRQVIQFIDLSLITKQAHNNMRGANFI
S-typhimurium	MSTAKLVKTKASNLLYTRNDVSESDKKATVELLNRRQVIQFIDLSLITKQAHNNMRGANFI
H-pylori	-----MKTFEILKHLQADAIVLFMKVHNHNNVRGTDFF
C-jejuni	-----MSTFKQLLQMQADAHLLWVFKHNNVRLGLQFF
L-lactis-A	-----MSNEHTQVELNQVADLSKASALVHOIHWWLRGPGFL
S-mutans	-MTNTITENIYASIIHQVEKKENSNGEKTAVLNQAVADLSKASIVHOVHWYMRGSGFL
L-monocytogenes	-----MKTINSVDTKETLNHQVANLNVTVKIHOIHWYMRGHNFF
L-innocua	-----MKTINSVDTKETLNHQVANLNVTVKIHOIHWYMRGHNFF
B-anthraxis-Dlp1	-----MNMKVIVELNKKQVADWSVLFTKLHNHFWYVKGPPQFF
B-anthraxis-Dlp2	-----MSTKTNVVEVLNKKQVANWVLYVKLHNYHWYVTPGPHFF
S-aureus	-----MSNQDQVVKELNKKQVANWTVAYTKLHNHFWYVKGPPNFF
L-lactis-B	-----MTKLTIDEKYAKELDKAEVDHMKPTAGAMLGHVLSNLFENVRLTGAGIY
E-coli	AVHEMLDGFRTALIDHLDTMAERAVQLGGV-----ALGTTQVINSKTPLKSYPLDIH-NV
S-typhimurium	AVHEMLDGFRTALTDHLDTMAERAVQLGGV-----ALGTTQVINSKTPLKSYPLDIH-NV
H-pylori	NVHKATEEIEYEGFADMFDDLAERIAQLGHH-----PLVTLSEALKLTRVKEETKTS-FHS
C-jejuni	SIHEYTEKAYEEMAEFLDSCAERVLQLGKE-----AITCQKVLMEAKSPKVAKDC-FTP
L-lactis-A	YLHPKMDLKDQDLDEHLEDEFAERLITIGGS-----PVSTLAEFDKNSKEMTPAVWVGKSN
S-mutans	YLHPKMDLMDLNGHLLDEISERLITIGGA-----PFSTLKEFDENSRLTEVTGTDWKSII
L-monocytogenes	TLHEKMDLYSEFGEQMDDEVAERLLAIGGS-----PFSTLKEFLENASVEEAPYTKPKTM
L-innocua	TLHEKMDLYSEFGEQMDDEVAERLLAIGGS-----PFSTLKEFLENASVEEAPYTKPKTM
B-anthraxis-Dlp1	TLHEKFEELYTESATHIDEIAERLLAIGGK-----PVATMKLEYLSSSQEAAYGE--TA
B-anthraxis-Dlp2	TLHEKFEELYNAGTYIDEIAERLLALEGK-----PLATMKLEYLSSSVNNEGSKS--SA
S-aureus	SLHVKFEELYNEASQYVDELAERLLAVGGN-----PVGTLTECLEGSIVKEAAKGY--SA
L-lactis-B	AKSPVKCELYLREIAKKEDEYFFKISDLLLLDENEIVPSTTEELFKYHKFTTEDPKAKYWD
E-coli	QDHLKELADRYAIVANDVRKAIGEAKE---DDTDADILTAASRDLDKFLWFIESNIE----
S-typhimurium	QDHLKELADRYAVVANDVRKAIGEAKE---DETDADILTAASRDLDKFLWFIESNIE----
H-pylori	KDIFKEILEDYKHKLEKFKELSNIAEKGDKVTVTYADDQLAKLQKSIWMLQAHLA---
C-jejuni	LEVIELIKQDYEYLLAEFKKLNEAAEKESDTTTAAFAQENIAKYKSLIWMIGATLGQACK
L-lactis-A	SERVKELIVAYKYLTQLFKDGIKAGDDGDVTVDTLYTTALGDIEKTIWMIIEAEVG---
S-mutans	TDHLKRLVQYVDYLSLSYQVGLDVTDEEDDAVSNDIFTAAQTEAQKTWMLQAEFGQAPG
L-monocytogenes	DQLMEDLVGTLELLRDEYQGGIELTDKEGDNVTNDMLIAFKASIDKHIMMFKAFGLKAPL
L-innocua	DQLMEDLVGTLELLRDEYQGGIELTDKEGDNVTNDMLIAFKASIDKHIMMFKAFGLKAPL
B-anthraxis-Dlp1	EEGMVQTLVNDYSALIQELKEGMEVAGEAGDATSADMLLAHTTLEQHWMLSLAFLK----
B-anthraxis-Dlp2	EEGMVEELSQDFNISKQLENAIEIAGNAGDDVSEDMFIMGQTSVDKHNMMFKSYLS----
S-aureus	EALLESFIDAFQAQNMFITRAIKLANKEEKFALAAVVVELLYGYNLQVIRHLAGDLGKAVA
L-lactis-B	
E-coli	-----
S-typhimurium	-----
H-pylori	-----
C-jejuni	M-----
L-lactis-A	-----
S-mutans	L-----
L-monocytogenes	E-----
L-innocua	E-----
B-anthraxis-Dlp1	-----
B-anthraxis-Dlp2	-----
S-aureus	-----
L-lactis-B	DFHDEDEDNDN

Fig. 9.1 Alignment of Dps family proteins. Dps proteins were aligned using ClustalW (<http://align.genome.jp>) and the Dps proteins of *E. coli* (Accession number AAD28292), *L. lactis* DpsA and DpsB (Accession number NP_268182 YP_001174737), *H. pylori* NAP (Accession number BAA96880), *S. aureus* MrgA (Accession number ZP_02760679), *B. anthracis* Dlp-1 and Dlp-2 (Accession number AAM18635 AAM18636), *L. monocytogenes* Fri (Accession number NP_464468), *L. innocua* Fri (Accession number CAC96173), *S. mutans* Dpr BAA96472), *S. enterica* Typhimurium Dps (Accession number NP_459808), *C. jejuni* Dps (Accession number YP_002344906)

9.3 Dps-DNA Co-Crystallization

When the DNA binding properties of Dps were initially studied it was noted that when *E. coli* Dps was heat treated to 65°C it lost its DNA binding ability while the DNA-bound Dps could be heated to 100°C without losing its binding properties (Almiron et al. 1992). These data suggested that in the bound state, Dps-DNA forms an extremely stable complex. The explanation for this stability was provided

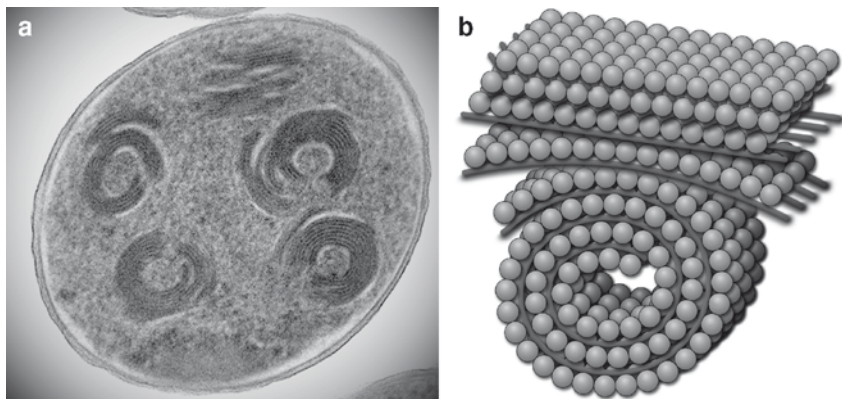


Fig. 9.2 Dps-DNA structures in stationary phase *E. coli* cells. **(a)** Electron micrographs of Dps-DNA toroidal structures in *E. coli* cells 24 h after onset of stationary phase. **(b)** Model for how the dodecamer Dps particles (*spheres*) and DNA (*lines*) assemble into toroidal structures that form the template for the Dps-DNA co-crystals (pictures kindly provided by A. Minsky)

by electron micrographs that revealed a highly ordered honeycomb-like lattice structure composed of ring-shaped, toroidal chromatin structures condensed by Dps as shown in Fig. 9.2 (Almiron et al. 1992; Wolf et al. 1999; Frenkiel-Krispin et al. 2004a). For the *E. coli* Dps it also became clear that prolonged incubation of purified Dps allowed it to enter a crystalline state and that addition of DNA greatly accelerated the crystallization process although the resulting crystals did not differ from those formed in the absence of DNA (Wolf et al. 1999). More detailed studies revealed gathering of every three neighboring dodecamers of Dps within a layer that results in a threefold symmetry surrounding each dodecamer with three larger holes and three smaller holes. Upon formation of the multi-layered structure in the crystal one layer of dodecamers is added on top of another layer with a shifted lateral position so that larger holes in one layer fit on top of a dodecamer in the neighboring layer. Several models have been proposed for the binding of DNA but most data are consistent with a structure in which Dps and DNA form alternate layers and the double stranded DNA helix is fitted within grooves formed between layers of Dps dodecamers (Wolf et al. 1999; Ren et al. 2003).

The unusual ability of Dps to organize DNA in higher-order crystal-like structures spurred a number of studies to examine if similar structures could be observed *in vivo*. Indeed, crystallization was readily observed *in vivo* in *E. coli* cells that either were genetically modified to overproduce Dps protein or following induction of Dps expression in starved *E. coli* cells. The structures resembled those seen *in vitro* containing both DNA and Dps and they were not observed in mutants lacking Dps (Wolf et al. 1999). Conversely Dps was the dominant protein associated with DNA when DNA was isolated from cells over-expressing Dps (Wolf et al. 1999). In the work of Frenkiel-Krispin et al. (2004a) DNA-Dps co-crystals were

observed after 2 days of starvation in *E. coli* cells over-expressing Dps and, although much less frequently, also in wild-type cells in stationary phase where Dps protein expression naturally is induced (see below). When examining Dps over-expression in *E. coli* cells in the early stage of starvation (24 h) the interaction between Dps and DNA resulted in toroidal structures that adopted different orientations within the cytoplasm. After 36 h of starvation tightly packed three-dimensional crystals started to appear while the ring shaped structures were still visible. At 48 h the cell is completely filled with packed crystal structures (Frenkiel-Krispin et al. 2004a). Detailed inspections of the electron micrographs revealed that the initial Dps-DNA toroidal structures appear to function as a nucleation site for the subsequent growth of the Dps-DNA co-crystal planes and depending on the location of the initial ring shaped structures the crystals can adopt different orientations within the cell (Frenkiel-Krispin et al. 2004a).

Curiously, overproduction of Dps protein in the exponential growth phase is not enough to induce Dps-DNA co-crystallization suggesting that other cellular or environmental factors influence the process (Frenkiel-Krispin et al. 2001). One likely explanation of this finding came from in vitro studies where the Dps mediated protection of DNA from nucleases was abolished by high concentrations of divalent cations while a reduction allowed Dps-DNA interactions and protection from digestion (Frenkiel-Krispin et al. 2001). Thus, in the presence of high concentrations of cations an electrostatic repulsion may keep Dps from binding to DNA and therefore the condensed Dps-DNA complex is not formed even in the presence of high concentrations of Dps (Frenkiel-Krispin et al. 2001). As discussed below a high concentration of cations is likely to be present in actively growing cells whereas the concentration falls as cells enter stationary phase. By this mechanism the DNA can enter a crystal-like state in a process where neither the process nor the signaling requires energy (in the form of ATP; the process is itself is likely to be energetically favorable), something that will be limiting during starvation in stationary phase. Also, when Dps is overproduced in response to oxidative stress the co-crystallization of Dps/DNA is not observed. This phenomenon may be related to the cation concentration as mentioned but the activity of another nucleoid-associated protein, Fis, also contributes (Ohniwa et al. 2006). In response to oxidative stress, Fis represses expression of both DNA gyrase and DNA topoisomerase I (Topo I), the enzymes that maintain DNA superhelicity, and consequently it was speculated that a reduction in the dynamic state of the chromosomal DNA topology hampers the crystallization process under such conditions. This notion was confirmed by the observation of nucleoid condensation during oxidative stress in cells lacking Fis and in cells overproducing Topo I and gyrase (Ohniwa et al. 2006).

The structural consequences on the nucleoid of Dps expression in stationary phase have also been investigated in organisms other than *E. coli*. In contrast to *E. coli* the *S. aureus* nucleoid does not aggregate upon entry to stationary phase, a finding that is paralleled by constant expression of the Dps homologue, MrgA, in both the exponential and the stationary growth phase (Morikawa et al. 2006). However, when cells are exposed to oxidative stress a massive production of MrgA is initiated and it is accompanied by a condensation of the bacterial chromosome,

a phenomenon not observed in cells lacking the *mrgA* gene (Morikawa et al. 2007). Toroidal chromatin structures are not uniquely associated with Dps binding as they have also been observed in other bacterial systems, such as in the early steps of *Bacillus subtilis* spore formation (Pogliano et al. 1995) and in association with the small, acid soluble DNA-binding proteins, SASP, expressed during sporulation (Ragkousi et al. 2000). For the *B. subtilis* spore the dense protein-DNA structures correlate with enhanced resistance (Frenkiel-Krispin et al. 2004b). Thus, chromosome condensation as a protective measure is conserved across bacterial species and biological systems.

9.4 Dps DNA Binding

Dps is a non-specific DNA binding protein that binds to both supercoiled and linear DNA and single-stranded RNA (Almiron et al. 1992; Wolf et al. 1999). The primary amino acid sequence of Dps does not reveal any conventional DNA binding motifs (Almiron et al. 1992). However, when the structure of the protein was examined it was proposed that the positive charge provided by lysine residues located in the highly mobile N-terminal region could provide DNA binding through the cooperative action of three neighboring Dps dodecamers (Grant et al. 1998). This finding was supported by the correlation between the number of positively charged residues in the N-terminal region and the ability to form large Dps-DNA complexes (Ceci et al. 2004). In addition to *E. coli* Dps, DNA binding has been demonstrated for *Bacillus subtilis* MrgA, *Synechococcus* DpsA and *Mycobacterium smegmatis* Dps (Chen and Helmann 1995; Gupta and Chatterji 2003; Gupta et al. 2002; Peña and Bullerjahn 1995). For Dps homologues in *Deinococcus radiodurans* (Dps-1) and *Lactococcus lactis* (DpsA and DpsB) the N-terminal region is also involved in DNA binding although in these proteins it additionally involves an intra-subunit metal-binding site located in an alpha helix of the N-terminal region (Kim et al. 2006; Stillman et al. 2005; Bhattacharyya and Grove 2007). Surprisingly with this structure the requirement for positively-charged lysine residues is abolished as substitutions of the Lys 9,15 and 16 by Glu did not abolish DNA binding (Stillman et al. 2005). One of the *Mycobacterium smegmatis* paralogues, MsDps1, carries a C-terminal extension of which the most distal 16 amino acids are required for DNA binding (Roy et al. 2007).

From the experiments described above it was predicted that several Dps family members should be unable to bind DNA due to the lack of an unstructured N-terminal region. These include *L. innocua* and *Bacillus anthracis* Dlp-1 and Dlp-2 where the N-terminal region is very short and the *Agrobacterium tumefaciens* Dps carrying an inflexible N-terminal region (Bozzi et al. 1997; Ilari et al. 2000; Papinutto et al. 2002; Ceci et al. 2003). A non-flexible N-terminal region may also explain the inability of *Streptococcus mutans* Dpr to bind DNA (Yamamoto et al. 2002). In *Helicobacter pylori* a Dps homologue was identified termed Hp-Nap for *H. pylori* neutrophil activating protein that initially proved unable to bind DNA

(Tonello et al. 1999; Ceci et al. 2003) but upon expression in *E. coli* it co-localized with the chromosome (Cooksley et al. 2003). The explanation for this discrepancy was recently found as the protein is characterized by an overall positive surface charge that abolishes the need for a positive N-terminal region and allows HpNap-DNA interactions in a pH-dependent manner: At pH 8.5 the protein is largely unbound, at pH 7.5 the primary DNA binding occurs and at pH 6.5 or lower secondary DNA binding involves looping and condensation (Ceci et al. 2007). These findings suggest that the DNA binding ability of other “non-DNA binding” Dps homologues should be re-investigated and although many Dps proteins bind DNA the binding domain is not conserved between species.

9.5 Iron Binding and Ferroxidase Activity

Structurally and functionally the Dps family members resemble the ferritins in forming spherical shells of monomers, but while the ferritins consist of 24 subunits binding more than 4,000 iron ions (Harrison and Arosio 1996), members of the Dps protein family form a twelve-monomer spherical structure that can bind approximately 500 iron molecules (Bozzi et al. 1997; Grant et al. 1998; Ishikawa et al. 2003; Tonello et al. 1999; Yamamoto et al. 2002; Zhao et al. 2002). Each Dps monomer is composed of four major helices forming a bundle as is the case for the ferritins but additionally it contains a minor helix separating the major helices B and C and lacks a fifth helix present in the ferritins. The ability to bind and detoxify Fe^{2+} is associated with ferroxidase activity that catalyzes the oxidation of Fe^{2+} to Fe^{3+} (Zhao et al. 2002) and is generally found well conserved among the Dps family members. In *Listeria innocua* the iron binding site was demonstrated to involve the Glu62 and Asp58 located on the B helix of one subunit, and on another subunit the A helix, His31 and through a water molecule forming a hydrogen bond with the His43 with additional stabilization by Asp47 (Ilari et al. 2000). Thus it was a striking finding that in contrast to the ferritins, the ferroxidase centre of Dps is located at the interface of two individual Dps subunits (Ilari et al. 2000, 2002; Zhao et al. 2002; Ren et al. 2003). Another notable difference is that the oxidation performed by Dps proteins is most effectively accomplished by H_2O_2 in contrast to the ferritins that use O_2 (Zhao et al. 2002). Interestingly, Dps proteins are also able to bind zinc and the ferroxidase activity can in fact be inhibited in the presence of zinc and terbium (Yamamoto et al. 2002; Stefanini et al. 1999). For the Dps allele of *Streptococcus suis* binding of zinc and terbium almost completely abolished iron binding through interaction with the ferroxidase centre and a novel zinc binding site was identified involving the highly conserved His44 and less conserved His40 residues (Havukainen et al. 2008). These findings suggest that Dps proteins may in general play a yet unidentified role in detoxification of zinc. Notable exceptions are DpsB from *Lactococcus lactis* and Dps1 from *B. anthracis* (Stillman et al. 2005; Liu et al. 2006). The two *L. lactis* Dps homologues, DpsA and DpsB are characterized by lacking conventional ferroxidase centres but instead contain N-terminal intrasubunit metal binding site (Stillman et al. 2005).

9.6 Control of Dps Expression

While studies of gene expression may reveal interesting regulatory mechanisms they also provide an indication of when specific proteins are required for performing certain cellular functions. In *E. coli* it was noted early on that in exponentially growing cells no *E. coli* Dps protein was detectable whereas in 1-day old cell cultures expression had increased to 20,000 monomers of Dps per cell thus demonstrating that stationary phase strongly induces Dps expression (Almiron et al. 1992). This induction was not seen in cells lacking the stationary phase sigma factor, σ^S thus strongly suggesting that σ^S is needed for transcription of *dps* in this growth phase (Almiron et al. 1992). In this growth phase the induction is also stimulated by another protein, integration host factor (IHF), that is a nucleoid-associated protein, whose cellular concentrations increases in stationary phase (Altuvia et al. 1994). In exponentially growing *E. coli* cells the *dps* gene is expressed by the σ^{70} sigma factor in an OxyR dependent manner (Altuvia et al. 1994). OxyR is a peroxide sensory protein that directly senses the oxidative state of the cell through the redox state of its cysteine residues (Zheng et al. 1998). In the reduced state, OxyR binds to two adjacent major grooves separated by one helical turn, while during peroxide stress the cysteine residues Cys-199 and Cys-208 form an intramolecular disulfide bond leading to binding of four adjacent major groove regions and activation of transcription by direct contact with RNA polymerase (Lee et al. 2004). Although the OxyR protein is present in stationary phase the OxyR-dependent induction occurs only during exponential growth suggesting that other factors contribute to the regulation of Dps expression under oxidative conditions in this growth phase (Altuvia et al. 1994). The OxyR-dependent *dps* gene regulation is also conserved in other Gram-negative bacteria distantly related to *E. coli* such as in the orthodontal pathogen, *Porphyromonas gingivalis* (Ueshima et al. 2003).

In the initial characterization of *dps* transcription in *E. coli* by σ^S and σ^{70} it was noted that the transcriptional start site was the same for both sigma factors and until recently the mechanism responsible for sigma factor selection has been obscure. However, in an elegant study by Grainger et al. (2008) it is demonstrated that both Fis and another nucleoid-associated protein, H-NS (Dorman 2004; Tendeng and Bertin 2003; Atlung and Ingmer 1997) are able to bind to the *dps* promoter region and repress expression. Both Fis and H-NS selectively prevent transcription initiated from the σ^{70} promoter whereas activity from the σ^S recognized promoter is unaffected. Fis acts by trapping the RNA polymerase containing σ^{70} at the promoter whereas H-NS displaces the σ^{70} -containing RNA polymerase but not the σ^S containing polymerase (Grainger et al. 2008). The authors propose a model in which the *dps* σ^{70} promoter is repressed during exponential growth by the combined action of Fis and H-NS but as cell density increases and the Fis concentration decreases, the Fis trapped σ^{70} RNA polymerase is released allowing access for the σ^S RNA polymerase (Grainger et al. 2008). Induction during exponential growth by e.g. oxidative stress is not accounted for in the model but it can be envisioned to occur through an OxyR dependent release of the Fis trapped σ^{70} containing RNA

polymerase or by an OxyR mediated change in the structure of the Fis-trapped σ^{70} complex that allows open complex formation and initiation by the σ^{70} RNA polymerase (Schnetz 2008).

Expression of *dps* in *E. coli* is stimulated by a number of factors such as carbon, nitrogen and phosphorous starvation as well as by osmotic stress and low temperature (Matin 1991; Lomovskaya et al. 1994; Phadtare and Inouye 2004). Using reporter gene fusions the induction of *dps* expression by carbohydrate starvation was found to occur at the transcriptional level whereas nitrogen starvation increased the amount of Dps protein tenfold without a significant increase in transcription (Lomovskaya et al. 1994). The posttranscriptional increase in amount of Dps during nitrogen starvation could be mediated by increased stability of Dps as the protein is post-translationally controlled by proteolysis (see below). The induction of *dps* expression was observed when shifting cells from 37°C to either 15°C or 23°C and it was eliminated in the *cspA cspB cspG cspE* quadruple-deletion strain as well as in a mutant lacking the stationary phase sigma factor, RpoS suggesting that both σ^S and the cold shock proteins are involved in expression at low temperature (Phadtare and Inouye 2004). In *S. Typhimurium* the *dps* promoter region carries binding sites for IHF and OxyR and in addition two DNA binding sites for the ferric uptake regulatory protein, Fur (Yoo et al. 2007). Fur is an iron-sensing regulator that controls the expression of genes for siderophore biosynthesis and iron transport and although Fur is commonly thought of as a metal-dependent repressor, Fur also activates the expression of many genes by either indirect or direct mechanisms (Lee and Helmann 2007). However the role of Fur in the control of *S. Typhimurium* *dps* expression is controversial. Velayudhan et al (2007) reported that expression of *dps* is strongly derepressed in the absence of Fur particularly in iron-replete medium, whereas Yoo et al (2007) showed that Fur substantially stimulates *dps* expression irrespectively of the iron status. The reason for these discrepancies is unknown but they may point to a dual role of Fur in both repressing and stimulating *dps* expression possibly depending on strain background and growth medium composition.

In Gram-negative bacteria distantly related to *E. coli* the expression of *dps* is more variably regulated. In the human pathogen *Helicobacter pylori* the expression of HP-NAP is induced in stationary phase (Cooksley et al. 2003) but also by acidic conditions possibly reflecting the adaptation of the organism to the environment during gastric colonization (Wen et al. 2003). *Campylobacter jejuni* is also a human pathogen closely related to *H. pylori* but in this organism the Dps protein is constitutively expressed at a level detectable by Western blot analysis in both exponential and stationary phase and is surprisingly not induced by oxidative stress (Ishikawa et al. 2003). *C. jejuni* is characterized by expressing only two alternative sigma factors and by the presence of regulatory proteins resembling those normally encountered in Gram-positive bacteria such as the peroxide regulator, PerR, that also in this organism responds to oxidative stress (Parkhill et al. 2000; van Vliet et al. 1999). The lack of oxidative stress induction of *C. jejuni* Dps expression suggests that the protein either is regulated at the post translational level or that the Dps is not required during oxidative stress in this organism.

In Gram-positive bacteria the expression of Dps homologues is also regulated by adverse conditions and maybe to an even greater degree than in Gram-negative

bacteria (see below). Generally Dps expression is induced by entry into the stationary growth with the MrgA homologue in *Staphylococcus aureus* being a marked exception (Morikawa et al. 2006; Chen et al. 1993; Polidoro et al. 2002). In several *Bacillus* spp. the chromosome encodes two Dps homologues suggesting that the biological activities provided by single Dps homologues in most organisms are split between two alleles in Bacilli thus providing an interesting opportunity to determine the differential regulation associated with the individual properties of the two alleles (Chen et al. 1993; Liu et al. 2006). In *B. subtilis* one protein, MrgC (DpsA, Dps2) is induced by entry into stationary phase, by glucose starvation, heat and osmotic stress through the activity of the stationary phase sigma factor σ^B and at the transcriptional level expression is stimulated by iron limiting conditions in mid-exponential growth, heat, salt and ethanol stress (Chen et al. 1993; Antelmann et al. 1997). In contrast, expression of the other Dps homologue, MrgA is induced by peroxidative stress and entry into stationary phase and the expression is strongly repressed by several different metal ions such as manganese and iron (Antelmann et al. 1997; Chen et al. 1993, 1995; Chen and Helmann 1995). Thus, there is a common but differential regulation of the two *dps* homologues in response to metal ions where *mrgC* is repressed by iron and *mrgA* is repressed in response to multiple ions (Chen et al. 1993). This control is mediated by two different members of the Fur family of regulators, namely Fur that controls the expression of *mrgC* and PerR that regulates the expression of *mrgA* (Chen et al. 1993, 1995; Bsat et al. 1996, 1998). PerR is the major regulator of the inducible peroxide stress response and like Fur it carries a structural Zn(II) binding domain in addition to a regulatory metal binding domain that through binding of Fe(II) and Mn(II) activates the DNA binding. Interestingly the two forms of PerR (carrying either Fe(II) and Mn(II)) differ in their sensitivity to peroxide stress (reviewed by Lee and Helmann (2007)). The structural similarities between Fur and PerR are also reflected in the similar DNA sequences to which PerR and Fur binds (Herbig and Helmann 2001; Baichoo and Helmann 2002). In fact, mutations of just two base pairs of the PerR binding site in the *B. subtilis* *mrgA* promoter region allowed dual regulation by both PerR and Fur in vivo (Fuangthong and Helmann 2003).

In general the control of *dps* expression in Gram-positive bacteria seems to be subjected to equally complicated regulation and to be regulated by PerR (Yamamoto et al. 2000; Brenot et al. 2005; Olsen et al. 2005; Rea et al. 2005; Morikawa et al. 2006; Nicodème et al. 2004). Particularly detailed studies of gene expression have been performed in the two bacterial species of *Listeria*, namely *L. monocytogenes* and *L. innocua*. The expression of the *L. monocytogenes* homologue, Fri is induced by a number of stress conditions including heat and cold shock, and to a lesser extent SDS, ethanol and deoxycholate (Hébraud and Guzzo 2000; Liu et al. 2002; Phan-Thanh and Gormon 1995; Olsen et al. 2005; Mohamed et al. 2006). Northern blot analysis revealed that induction by heat and cold shock occurred at the transcriptional level and that the Fri protein is expressed from a monocistronic mRNA (Hébraud and Guzzo 2000). Controversial was the finding that transcription of *fri* is induced by stationary phase in both *L. monocytogenes* and *L. innocua* (Polidoro et al. 2002) when monitored by dot blot as recent data by the same group reveal an almost complete abolishment of *fri* transcription in stationary growth phase when

examined by Northern blot (Fiorini et al. 2008). The reason for this discrepancy remains unknown although the different experimental designs may be responsible. In *L. monocytogenes* the *fri* promoter region contains three promoters of which two are recognized by the general σ^{70} sigma factor, σ^A and one by the general stress sigma factor, σ^B (Olsen et al. 2005). In addition, *fri* expression is negatively regulated by both PerR and Fur that bind to overlapping sequences positioned just upstream of the *fri* ribosome binding site and are responsible for the induction of *fri* expression by iron-limiting conditions (Polidoro et al. 2002; Olsen et al. 2005; Fiorini et al. 2008). The possibility that both proteins bind to essentially the same DNA region indicates that as in *E. coli* there are multiple levels of regulation of *dps* transcription and that multiple signals may be integrated through a combination of promoter selection and derepression/re-repression by both Fur and PerR.

In addition to the transcriptional control of *dps* expression the cellular level of Dps is regulated at the post-transcriptional level. Interestingly, Stephani et al. (2003) found that the amount of Dps is kept under tight control by proteolysis. During exponential growth there is a rapid degradation of Dps protein but upon carbon starvation or oxidative stress the turnover of Dps is dramatically stabilized (Stephani et al. 2003). The protease responsible for the degradation in exponential phase is the ClpXP proteolytic complex that is composed of the substrate specificity unit, the ClpX ATPase, and the proteolytic core barrel composed of ClpP monomers (Frees et al. 2007). The positive effect of proteolysis on Dps expression seems to occur at least at two levels, firstly in the reduced turnover of the Dps protein and secondly in the stabilization and accumulation of σ^S that also is a target for ClpXP degradation (Stephani et al. 2003). While little is known of post-transcriptional regulation of Dps in other bacteria results from *L. monocytogenes* suggest that also in Gram-positive bacteria Dps is highly stable in the stationary phase as the protein remains detectable despite a dramatic decrease in transcription (Fiorini et al. 2008).

For a protein to perform its biological function optimally it needs to be timely expressed and to be active. For Dps we have only very little information to suggest that the activity of the protein is modulated through post-translational modifications, however, one report indicates that this may in fact be the case. In *Salmonella* a small fraction of the Dps protein was recently found to be glycosylated by residues including mannose and the glycosylated form of the protein was most prominent at the transition from exponential to stationary phase (Hanna et al. 2007, 2008). Thus, the activity of Dps may be modulated by glycosylation and such modification could be responsible for the growth phase dependent changes in structure and protein stability.

The discovery that Dps is a DNA binding protein led to the hypothesis that it was also involved in gene regulation. Almirón et al. (1992) showed by two-dimensional protein gel electrophoresis that the proteome of cells lacking Dps was dramatically different from that of wild-type cells in stationary phase after 3 days of starvation but not in exponential growth phase. The changes observed included both the appearance of proteins not expressed in wild-type cells as well as the disappearance of proteins present in stationary phase cultures of wild type cells. This phenomenon may be specific for

E. coli because in *B. subtilis* no major changes in the protein expression pattern were observed when wild type and *mrgA* mutant cells were compared (Antelmann et al. 1997). However, it has been shown more recently that inactivation of the gene encoding the *L. monocytogenes* Fri protein altered expression of 15 different proteins of which most were related to stress tolerance and one was listeriolysin, a primary virulence factor (Dussurget et al. 2005). One obvious mechanism for Dps-controlled changes in gene expression could involve the binding of the protein to promoter regions of the genes affected with a consequent modulation of transcription initiation. This would however involve some degree of sequence specificity in the binding, which has not yet been reported. Alternatively, it is possible that Dps binding constrains DNA either locally or globally and induces changes in superhelicity. DNA superhelicity has been reported to be a global sensor of the microbial habitat and to affect gene expression (Hatfield and Benham 2002). So far the influence of Dps on overall superhelicity has not been investigated but Chodavarapu et al (2008) noted that inactivation of *dps* did not affect the superhelicity of a plasmid. However, additional experiments are required to resolve this issue.

When comparing the regulation of Dps homologues in Gram-positive and Gram-negative bacteria there are some common features. In many organisms, expression of *dps* involves several promoters recognized by the σ^{70} sigma factor and the general stress sigma factors, RpoS or sigma B. In addition several regulatory proteins bind to the promoter region and in some cases there may even be a competition between regulators bound to the promoter region such as H-NS, Fis, OxyR in *E. coli* and Fur and PerR in *Listeria monocytogenes*. At the post-transcriptional level little is known yet, however some indications point to a stabilization of the Dps protein in stationary phase in several organisms, suggesting that proteolysis is important for the timely expression of Dps (Stephani et al. 2003; Fiorini et al. 2008). Thus, Dps expression is controlled at multiple levels in response to a variety of environmental stimuli.

9.7 Biological Impact of Dps

It has been recognized since the earliest studies of Dps family members that these proteins play a central role in protection against oxidative stress. Many reports have documented that in the absence of Dps bacterial cells are sensitive to peroxide stress such as that associated with H_2O_2 (Chen and Helmann 1995; Martinez and Kolter 1997; Ishikawa et al. 2003; Velayudhan et al. 2007; Ueshima et al. 2003; Halsey et al. 2004; Pulliainen et al. 2005; Olsen et al. 2005; Brenot et al. 2005). Generally, two modes of protection have been observed. One involves the ability of the Dps-like protein to bind non-specifically to DNA and to protect it physically from damage (Altuvia et al. 1994; Martinez and Kolter 1997; Wolf et al. 1999) and the second mode involves sequestration of iron and detoxification via the ferroxidase activity (Zhao et al. 2002). In vivo, the contribution of each of these two activities to the biological protection against oxidative stress can be difficult to assess and it

may also vary between organisms. The protective power of Dps as a DNA binding protein has most clearly been demonstrated by in vitro studies. Purified *E. coli* Dps protects DNA from nicking in an in vitro DNA damage assay while heat-inactivated Dps fails to do so (Martinez and Kolter 1997). Also degradation of DNA proceeds rapidly in the presence of H_2O_2 and Fe(II) whereas both the *L. innocua* and the *E. coli* Dps proteins almost abolish degradation (Su et al. 2005). For the *H. pylori* Hp-NAP protein gel retardation studies reveal that this protein protects DNA from DNases only at those lower pH values at which the protein is able to bind DNA (Ceci et al. 2007).

The protective role of Dps DNA binding is more difficult to assess in vivo. In mutants lacking Dps, macroscopic DNA damage, manifested as single-strand breaks and genomic degradation, was observed in stationary-phase cells exposed to H_2O_2 (Martinez and Kolter 1997; Yamamoto et al. 2004); the *dps* mutants were also sensitive to UV and gamma irradiation (Morikawa et al. 2006; Nair and Finkel 2004). In contrast, Dps does not offer protection against more specific DNA damaging agents in vivo such as mitomycin C or metronidazol suggesting that if DNA binding is occurring it is not protective against these agents (Ueshima et al. 2003). In a recent study of the *Salmonella* complement of ferritin and ferritin-like proteins, it was noted that removal of Dps did not change the free iron content appreciably. However, the cells became highly sensitive to H_2O_2 suggesting that the mode of Dps protection under these conditions is due chiefly to the physical binding of Dps to DNA (Velayudhan et al. 2007).

The activity of Dps may also change with the growth phase of cells. When exponential phase *E. coli* cells are treated with H_2O_2 both wild type and *dps* mutant cells rapidly lose viability. In contrast when 3-day starved cells are exposed to H_2O_2 the wild type is essentially unaffected while the *dps* mutant rapidly loses viability at high peroxide concentrations (Almiron et al. 1992; Nair and Finkel 2004). In stationary phase, with its associated massive production of Dps, chromosome condensation occurs and co-crystals of Dps and DNA can be observed (Martinez and Kolter 1997; Wolf et al. 1999). Moreover, Frenkiel-Krispin et al. (2001) noted that in actively growing cells the chromatin is randomly spread over the cytoplasm, an observation that is consistent with a state of high metabolic activity. In contrast, bacteria starved for 2 days exhibit de-mixing of chromatin and ribosomes, with the ribosomes located at the periphery of the cytoplasm. X-ray scattering measurements have revealed diffraction patterns consistent with DNA-Dps co-crystal structures in which Dps and DNA form stacked alternating layers. The protection against H_2O_2 observed in cells where the DNA has been compacted with Dps suggests that the direct interaction between the DNA and Dps is in fact protecting the cells. Factors that in the bound state may also reduce susceptibility to peroxide stress are the reduced overall nucleoid surface (Morikawa et al. 2006) and the structured frame of the DNA-Dps complex that through restricted diffusion may allow repair of DNA damages such as double strand breaks (Levin-Zaidman et al. 2003).

Interestingly, the compaction of DNA by Dps does not take place in exponentially growing cells that are overproducing Dps artificially suggesting that a large amount of Dps per se is not sufficient to induce the structural changes (Martinez and Kolter 1997). This finding is in marked contrast to other nucleoid-associated proteins, such

as H-NS and SASP, where overproduction leads to nucleoid condensation and loss of viability (Setlow et al. 1991; Spurio et al. 1992). As mentioned above Frenkiel-Krispin et al. (2001) investigated factors that affect the Dps-DNA co-crystallization process and found that the concentration of divalent cations are critical for the condensation to occur. Both low and high concentrations of doubly charged cations prevents co-crystallization and in vivo addition of 5 mM Mg^{2+} to the growth medium prevent de-mixing of chromatin and ribosomes in stationary phase. As the concentration of free Mg^{2+} ions within *E. coli* cells is essentially identical to that in the environment, the authors proposed that extracellular divalent cations provide a signal for the intracellular DNA-Dps co-crystallization. Thus, during active growth and with a plentiful supply of nutrients, the divalent cation concentrations will be high and electrostatic repulsion will prevent DNA-Dps complex formation. In contrast, during starvation the divalent cation concentration is gradually reduced and below a threshold value ion bridges can be formed between DNA and Dps, resulting in binding. When fresh nutrient including divalent cations are supplied, the reverse process occurs (Frenkiel-Krispin et al. 2001). During prolonged starvation the divalent cation concentrations will be reduced to a level that does not sustain the DNA-Dps ion bridges and this correlates with the observation that in cells starved for more than 6 days the separation of ribosomes from the chromatin occurs independently of Dps (Frenkiel-Krispin et al. 2001). Thus, here Dps functions as a cellular sensor that provides a non-inducible regulatory mechanism which in cells with minimal metabolic activity and very little available energy can induce the binding of DNA by Dps and consequently the associated physical protection.

Some bacterial Dps proteins have proven unable to bind DNA including *L. innocua* Fri, *S. mutans* Dpr and Dlp-1/Dlp-2 from *B. anthracis*. For these organisms the protective role of Dps is likely to occur through iron binding and ferroxidase activity (Bozzi et al. 1997; Papinutto et al. 2002; Tonello et al. 1999; Yamamoto et al. 2002; Pulliainen et al. 2005). As mentioned above the primary oxidant of the ferrioxi-dase centre of the *E. coli* Dps is H_2O_2 although ferroxidation can occur at a much lower rate with O_2 (Zhao et al. 2002). That the same is true in other organisms is suggested by findings that *dps* mutants often are sensitive to H_2O_2 but no more sensitive to oxygen than the wild type (Ueshima et al. 2003; Brenot et al. 2005). Interestingly, studies of the two Dps homologues from *B. anthracis*, Dpl1 (Dps1) and Dpl2 (Dps2) show that while both proteins react similarly with $O_2/Fe(II)$ they have different reactions with $H_2O_2/Fe(II)$. For the Dpl2 protein the ferroxidase activity is only three-fold higher with H_2O_2 as oxidant compared to the reaction with O_2 as oxidant whereas in *E. coli* this ratio is more than 100-fold. Further, the Dpl1 protein has no ferroxidase activity at all with H_2O_2 and the activity with O_2 as oxidant is comparable to Dpl2 (Liu et al. 2006) suggesting, that the two proteins may be differentially required in various environments. In contrast, in the lactic acid bacteria the protective effect of Dps seems equally important when exposed to oxygen and peroxides. Lactic acid bacteria lack cytochromes and other heme containing proteins, including catalase that catalyzes the decomposition of H_2O_2 to O_2 and H_2O , but they can still grow in the presence of oxygen (Higuchi et al. 2000). *Streptococci* belong to the lactic acid bacteria and in *S. mutans* the *dpr* mutant has a severely reduced plating efficiency

under aerobic conditions. This defect can be suppressed by addition of catalase or the iron chelating agent deferoxamine to the growth medium suggesting that the toxicity is caused by the conversion of intracellular H_2O_2 to hydroxyl radicals via free iron ions and the Fenton reaction (Yamamoto et al. 2000). Both in *S. suis* and *S. mutans* inactivation of the Dpr protein increases the free cellular iron content as shown by the susceptibility of a *dpr* mutant to the iron-requiring antibiotic, streptonigrin (Yamamoto et al. 2004; Pulliainen et al. 2005). Interestingly lactic acid bacteria generally only contain one type of ferritin-like protein, namely Dps and therefore, in these organisms Dps may have acquired the ability of the ferritins to use oxygen efficiently for the Fe(II) oxidation process. Yet studies of the related *L. innocua* Dps protein suggest that it still has a preference for H_2O_2 as oxidant and the authors propose that it may be linked to the histidine residues that are present in Dps proteins but not in the ferritins (Su et al. 2005). Additionally, in these organisms Dps seems to be particularly important for iron storage, possibly as Fe(III) following Fe(II) oxidation (Pulliainen et al. 2005), as growth is restricted by conditions where iron is limiting (Olsen et al. 2005; Dussurget et al. 2005; Mohamed et al. 2006), which is not observed for Dps proteins in organisms from outside the lactic acid bacteria (Ueshima et al. 2003; Velayudhan et al. 2007; Ishikawa et al. 2003). Thus, we and others have speculated that storage and solubilization of Fe(II) may in fact be a very important role of the Dps proteins in this group of bacteria, a notion that is supported by the finding that the amount of iron that is associated with Dps is highly dependent on the iron concentration in the growth medium (Polidoro et al. 2002; Olsen et al. 2005).

The protective role of Dps may also depend on the oligomeric state of the protein as in vitro studies of the *Mycobacterium smegmatis* Dps reveal a stepwise assembly of the generic dodecamer as trimers are first formed and after overnight incubation at 37°C the trimers assemble into the dodecamer. In this system trimers are unable to bind DNA but still protect against damage produced by a combination of iron and hydrogen peroxide (Gupta and Chatterji 2003).

Irrespective of the mode of Dps action, both direct DNA binding with protection against oxidative stress as well as the sequestration of iron are activities that could be very important in protecting the cells against mutations. However, the influence of Dps on mutation frequency has only been investigated in a few instances. By using a set of strains containing *lacZ* mutations that allow rapid detection of each of the six possible base substitutions Martinez and Kolter (1997) showed that in *E. coli* Dps can specifically protect against GC-to-TA transversions in stationary phase cultures treated with H_2O_2 and against other base substitutions upon expression in exponential phase. Also overexpression of Dps partially suppresses the mutator phenotypes associated with defects in the oxidative repair enzymes encoded by *mutY* and *mutM* (Martinez and Kolter 1997; Lu et al. 2001). Interestingly, Dps only offers protection against oxidative stress agents as no difference in mutation rate was observed after treatment with the alkylating agent ethyl methanesulfonate (Martinez and Kolter 1997). However, in another organism, *Pseudomonas putida*, the absence of *dps* does not affect the mutation frequency even in cells also lacking oxidative repair enzymes. This observation was made using a test system that monitors base substitutions by elimination of the TAG stop codon in the phenol monooxygenase gene, *pheA* (Saumaa et al. 2007).

In *H. pylori* inactivation of the *napA* gene encoding HP-NAP does not per se affect mutation frequencies but when it is combined with a mutation removing the alkylhydroperoxide reductase, AhpC, the mutation frequency increases (Olczak et al. 2002). These results indicate that Dps may reduce the mutation frequency but only under certain conditions and that the impact may vary among bacteria.

In addition to the well-studied phenotypes already mentioned, the inactivation of *dps* is also accompanied by many pleiotropic effects such as sensitivity to high temperature following a shift of stationary phase cultures from ambient to high temperature and sensitivity to alkaline and acidic pH that are not easily explained with the traditional view of Dps (Nair and Finkel 2004). One possibility is that the altered gene expression pattern observed in stationary phase cells lacking Dps (Almiron et al. 1992) may be responsible for such differences. Alternatively, Dps may harbour still undiscovered and new biological activities. That the latter may in fact be the case is indicated by the recent findings published by Chodavarapu and co-workers that revealed a completely new function of Dps in DNA replication (Chodavarapu et al. 2008). Very interestingly, Dps interacts with the N-terminal part of the replication initiator protein, DnaA and in an in vitro assay inhibits the DnaA-dependent unwinding at the origin of replication, *oriC* (Chodavarapu et al. 2008) which is the region of the chromosome that contains a high number of DnaA-binding sequences and where the DNA duplication process begins (Bramhill and Kornberg 1988). This notion was confirmed by in vivo experiments revealing a reduced frequency of initiation of new rounds of replication in the presence of Dps. To explain the results the authors propose a model in which elevated level of Dps as experienced during oxidative stress reduces new rounds of DNA replication by interacting with DnaA to inhibit its function in initiation while also protecting DNA from further oxidative damage (Chodavarapu et al. 2008). However, replication is not completely blocked by Dps and the residual replication allowed during the mutagenic conditions of, for example, oxidative stress may provide the cells with an opportunity to increase mutation frequency and consequently the genetic variation of the population leading to survival of those cells that carry favorable mutations (Chodavarapu et al. 2008). The link between control of DNA replication, protection of DNA against damage and oxidative stress intriguingly suggests that Dps may be central in allowing cells to respond to conditions of starvation and oxidative stress by play a pivotal role in deciding when cells are ready for adaptive mutations. Thus, additional research should be directed at revealing the role of Dps in mutagenesis and particularly in the occurrence of stress-induced mutagenesis that in an *rpoS*-dependent manner over time occurs in cells in stationary phase (Bjedov et al. 2003).

9.8 Dps Contribution to Virulence

Iron is essential for most bacterial species but for bacterial pathogens it plays a particularly important role. As part of the immediate innate immune defense, human and animal hosts have adopted various ways of limiting iron availability in

order to reduce pathogen proliferation. In the serum and extracellular fluids the iron is bound by transferrin whereas in the mucosal secretions lactoferrin sequesters the iron (Wooldridge and Williams 1993). In response, bacterial pathogens use iron as an important environmental stimulus in the control of virulence gene expression (Adams et al. 1990; Coulanges et al. 1996). The particular importance of iron for the virulence of pathogens has also been demonstrated for several pathogens. For example the intracellular pathogen, *Listeria monocytogenes* is more virulent in mice fed with excess iron compared to those fed on an iron-free diet and the effect can be reversed by addition of iron chelators (Sword 1966). In the case of mycobacterial infections it has been reported that increased dietary intake of iron is associated with a 3.5-fold increase in risk of developing pulmonary tuberculosis (Gangaidzo et al. 2001).

During the infection process bacteria are not only faced with the problem of low availability of free iron, but also with severe oxidative stress due to the release of reactive oxygen intermediates from phagocytes. When a phagocyte encounters a microorganism, the latter is surrounded by a portion of the phagocytic membrane, which invaginates forming a discrete phagosome (Hurst and Barrette 1989). Subsequently, oxygen consumption by the phagocyte is increased and a complex biochemical signaling system is initiated which activates a membrane associated NADPH-dependent oxidase. The activity of the oxidase converts oxygen to superoxide that can dismute to hydrogen peroxide and these compounds are released into the phagosome (Clark 1990). If a small amount of iron is present it catalyzes the conversion to hydroxyl radicals as part of the Fenton reaction. Thus, in the phagosome the environmental conditions are exactly those that stimulate the synthesis of Dps proteins and require the protective ability of Dps against oxidative and peroxide stress as well as storage and detoxification of iron. Moreover, in some types of infections phagosome maturation is accompanied by a decrease in Mg^{2+} concentration as part of the defense against pathogens (Eriksson et al. 2003) and at least in *E. coli* this signal is a trigger for the Dps-DNA bio-crystallization process to initiate and provide a further protection of the bacterial DNA against the hostile phagosome environment.

Helicobacter pylori was one of the first pathogens in which Dps was investigated (Evans et al. 1995). It is an unusual organism that is able to colonize the highly acidic environment of the gastric mucosa. Infections with *H. pylori* are characterized by chronic gastritis, peptic ulcer and gastric cancer (Dunn et al. 1997). During the infection *H. pylori* elicits an inflammatory cell response, and the severity of mucosal injury is correlated with the extent of neutrophil infiltration (Davies et al. 1994; Mizuki et al. 2000). As part of a study of *H. pylori* proteins that activate neutrophils, Yoshida et al. (1993) found a protein that in water extracts of *H. pylori* promoted neutrophil adhesion to endothelial cells. Later, this protein was identified as a Dps homologue termed HP-NAP and was shown to stimulate production of reactive oxygen species in neutrophils in a process possibly involving myeloperoxidase (Evans et al. 1995; Wang et al. 2008). More detailed studies revealed that binding to carbohydrate-containing macromolecules may be important in the host as HP-NAP binds to mucin, a highly glycosylated protein forming the main constituent of

mucus (Namavar et al. 1998), and to compounds of the acid glycosphingolipid fraction of neutrophils (Teneberg et al. 1997). In addition, HP-NAP is chemotactic for human neutrophils and monocytes and thus plays a role in recruiting these cells to the site of infection (Brisslert et al. 2005). Uniquely, HP-NAP seems to have a dual role in *H. pylori* pathogenesis as it both induces production of reactive oxygen radicals through neutrophil activation and at the same time is required for survival during oxidative and acidic stress conditions where DNA condensation mediated by HP-NAP occurs at low pH suggesting that this particular ability is needed during gastric colonization (Cooksley et al. 2003; Ceci et al. 2007). Interestingly, several of the mentioned functions of HP-NAP during infection must require a direct interaction between the HP-NAP protein and the host suggesting that during the infection process a substantial number of *H. pylori* cells must undergo spontaneous or programmed lysis. The bi-functionality of the HP-NAP protein was elegantly demonstrated by Wang et al. (2008) who observed that *napA* (encoding HP-NAP) mutant cells colonized mice equally as well as wild-type *H. pylori* cells but if the mice were first stimulated by heat inactivated *H. pylori* and then exposed to either wild-type or HP-NAP-deficient cells the colonization by mutant cells was greatly reduced. Conversely, if mice were stimulated by inactivated HP-NAP-lacking cells both mutant and wild type colonized equally well. Thus, both the induction of the hostile reactive oxygen species and the subsequent bacterial survival is dependent on HP-NAP. Later, other Dps proteins have been tested but neither *L. monocytogenes* Fri nor *B. anthracis* Dlp1 or Dlp2 were able to activate neutrophils (Zanotti et al. 2002). However, in *Streptococcus suis*, an important pig pathogen the Dpr homologue seems to function as an adhesin by binding to glycoprotein containing host cell surface elements and thus in part have similar properties as Dpr in *H. pylori* (Tikkanen et al. 1995; Pulliainen et al. 2003).

Based on the obvious link between oxidative stress resistance, iron detoxification and virulence, the role of Dps proteins has also been investigated in other bacterial pathogens. In *Salmonella enterica* sv. Typhimurium gene expression was examined in various host environments and expression of *dps* is upregulated in macrophages (Valdivia and Falkow 1996). The ability of *Salmonella* to survive and replicate within host phagocytes is absolutely essential for virulence and the reactive oxygen radicals generated by the NADPH phagocyte oxidase are needed to combat the bacteria (Vazquez-Torres and Fang 2001). In this process Dps seems to play an important part as *dps* mutant cells are impaired in survival in primary peritoneal macrophages and are less able than wild-type cells to kill mice infected intraperitoneally with either wild-type or *dps* mutant bacteria (Halsey et al. 2004). Also, fewer *dps* mutants can be recovered from liver or spleen following an intraperitoneal infection (Halsey et al. 2004; Velayudhan et al. 2007). Thus, in *S. Typhimurium*, Dps is an essential virulence factor.

Among the Gram-positive pathogens, *Listeria monocytogenes* is a serious food-borne pathogen of humans that can give rise to an array of symptoms ranging from a mild 'flu-like condition to meningitis and septicemia with death as a likely consequence (Kathariou 2002). *L. monocytogenes* also contains a homologue of Dps that has many names: Flp (Hébraud and Guzzo 2000), Fri (Polidoro et al. 2002;

Olsen et al. 2005) and Frm (Mohamed et al. 2006). We and others showed that Fri improves survival of cells exposed to peroxide stress or in stationary phase and it supports growth under iron-limiting conditions (Olsen et al. 2005; Mohamed et al. 2006; Dussurget et al. 2005). Also at high iron concentrations growth of the *fri* mutant is compromised possibly as a consequence of the toxic effect of the iron in promoting production of reactive oxygen species (Dussurget et al. 2005). The growth limitation of *fri* mutant cells in iron-deprived medium lead us to propose that Fri may be particularly important as an iron storage protein as *L. monocytogenes* does not express iron scavenging siderophores and thus relies on other means of acquiring and keeping the available iron (Olsen et al. 2005; Cowart and Foster 1985). The importance of the Fri protein for virulence has been investigated using cell culture and mouse infection models. Adhesion and invasion have been investigated both in macrophages and in HeLa epithelial cells and in both instances *fri* mutant cells invade the cells equally well as wild-type cells but intracellular survival is decreased for mutant cells (Olsen et al. 2005; Mohamed et al. 2006). When virulence is examined in mice injected intraperitoneally with bacteria in stationary phase there is a significant reduction in the number of mutant cells recovered compared to wild-type in spleen and liver whereas no difference in virulence was observed between wild-type and *fri* mutant cells when injected intravenously or when in exponential growth phase (Olsen et al. 2005; Mohamed et al. 2006; Dussurget et al. 2005). Thus, Dps offers protection particularly to cells in the stationary phase and only in some steps of the infectious cycle (Olsen et al. 2005).

The importance of Dps for virulence has also been investigated for more distantly related organisms. In the Lyme disease agent *Borrelia burgdorferi*, it was recently shown that the Dps homologue is required for the spirochaete to persist in the tick and to be transmitted to new hosts (Li et al. 2007). Also in the dental plaque organism *Porphyromonas gingivalis*, Dps is required for persistence and intracellular replication (Ueshima et al. 2003). Thus, there are many examples of pathogens where Dps proteins play an important part in the virulence of the organism. However, for some groups of pathogens like those belonging to the *Streptococci* there have not been any reports of Dps contributing to virulence although the proteins have been studied extensively (see above). For this group of bacteria the main role of the Dps protein seems to be related to the growth and survival of the organisms outside the host and the contribution of Dps to virulence is likely to be dependent on the particular infectious route selected by the pathogen. This notion is supported by the fact that some important pathogens like *Mycobacterium leprae*, *M. tuberculosis*, *Neisseria meningitis* and *Chlamydia pneumoniae* do not possess Dps-like proteins (Roy et al. 2008) further stressing that Dps proteins have been acquired or have evolved with particular attention to environmental challenges faced by the individual organisms.

Acknowledgements I am deeply grateful to Prof. Abraham Minsky for providing the micrographs displayed in Fig. 9.2 and for the helpful comments on the manuscript.

References

- Adams TJ, Vartivarian S, Cowart RE (1990) Iron acquisition systems of *Listeria monocytogenes*. Infect Immun 58:2715–2718
- Almirón M, Link AJ, Furlong D, Kolter R (1992) A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. Genes Dev 6:2646–2654
- Altuvia S, Almirón M, Huisman G, Kolter R, Storz G (1994) The *dps* promoter is activated by OxyR during growth and by IHF and σ^S in stationary phase. Mol Microbiol 13:265–272
- Andrews SC (1998) Iron storage in bacteria. Adv Microb Physiol 40:281–351
- Antelmann H, Engelmann S, Schmid R, Sorokin A, Lapidus A, Hecker M (1997) Expression of a stress- and starvation-induced *dps/pexB*-homologous gene is controlled by the alternative sigma factor σ^B in *Bacillus subtilis*. J Bacteriol 179:7251–7256
- Atlung T, Ingmer H (1997) H-NS: a modulator of environmentally regulated gene expression. Mol Microbiol 24:7–17
- Baichoo N, Helmann JD (2002) Recognition of DNA by Fur: a reinterpretation of the Fur box consensus sequence. J Bacteriol 184:5826–5832
- Bhattacharyya G, Grove A (2007) The N-terminal extensions of *Deinococcus radiodurans* Dps-1 mediate DNA major groove interactions as well as assembly of the dodecamer. J Biol Chem 282:11921–11930
- Bjedov I, Tenaillon O, Gérard B, Souza V, Denamur E, Radman M, Taddei F, Matic I (2003) Stress-induced mutagenesis in bacteria. Science 300:1404–1409
- Bozzi M, Mignogna G, Stefanini S, Barra D, Longhi C, Valenti P, Chiancone E (1997) A novel non-heme iron-binding ferritin related to the DNA-binding proteins of the Dps family in *Listeria innocua*. J Biol Chem 272:3259–3265
- Bramhill D, Kornberg A (1988) A model for initiation at origins of DNA replication. Cell 54:915–918
- Brenot A, King KY, Caparon MG (2005) The PerR regulon in peroxide resistance and virulence of *Streptococcus pyogenes*. Mol Microbiol 55:221–234
- Brisslert M, Enarsson K, Lundin S, Karlsson A, Kusters JG, Svennerholm AM, Backert S, Quiding-Järbrink M (2005) *Helicobacter pylori* induce neutrophil transendothelial migration: role of the bacterial HP-NAP. FEMS Microbiol Lett 249:95–103
- Bsat N, Chen L, Helmann JD (1996) Mutation of the *Bacillus subtilis* alkyl hydroperoxide reductase (*ahpCF*) operon reveals compensatory interactions among hydrogen peroxide stress genes. J Bacteriol 178:6579–6586
- Bsat N, Herbig A, Casillas-Martinez L, Setlow P, Helmann JD (1998) *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. Mol Microbiol 29:189–198
- Ceci P, Ilari A, Falvo E, Chiancone E (2003) The Dps protein of *Agrobacterium tumefaciens* does not bind to DNA, but protects it towards oxidative cleavage. X-ray crystal structure, iron binding and hydroxyl-radical scavenging properties. J Biol Chem 278:20319–20326
- Ceci P, Cellai S, Falvo E, Rivetti C, Rossi GL, Chiancone E (2004) DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus. Nucleic Acids Res 32:5935–5944
- Ceci P, Mangiarotti L, Rivetti C, Chiancone E (2007) The neutrophil-activating Dps protein of *Helicobacter pylori*, HP-NAP, adopts a mechanism different from *Escherichia coli* Dps to bind and condense DNA. Nucleic Acids Res 35:2247–2256
- Chen L, Helmann JD (1995) *Bacillus subtilis* MrgA is a Dps (PexB) homologue: evidence for metalloregulation of an oxidative-stress gene. Mol Microbiol 18:295–300
- Chen L, James LP, Helmann JD (1993) Metalloregulation in *Bacillus subtilis*: isolation and characterization of two genes differentially repressed by metal ions. J Bacteriol 175:5428–5437
- Chen L, Keramati L, Helmann JD (1995) Coordinate regulation of *Bacillus subtilis* peroxide stress genes by hydrogen peroxide and metal ions. Proc Natl Acad Sci USA 92:8190–8194

- Chodavarapu S, Gomez R, Vicente M, Kaguni JM (2008) *Escherichia coli* Dps interacts with DnaA protein to impede initiation: a model of adaptive mutation. *Mol Microbiol* 67:1331–1334
- Clark RA (1990) The human neutrophil respiratory burst oxidase. *J Infect Dis* 161:1140–1147
- Cooksley C, Jenks PJ, Green A, Cockayne A, Logan RP, Hardie KR (2003) NapA protects *Helicobacter pylori* from oxidative stress damage, and its production is influenced by the ferric uptake regulator. *J Med Microbiol* 52:461–469
- Coulanges V, André P, Vidon DJM (1996) Esculetin antagonizes iron-chelating agents and increases the virulence of *Listeria monocytogenes*. *Res Microbiol* 147:677–685
- Cowart RE, Foster BG (1985) Differential effects of iron on the growth of *Listeria monocytogenes*: minimum requirements and mechanism of acquisition. *J Infect Dis* 151:721–730
- Crosa JH (1997) Signal transduction and transcriptional and posttranscriptional control of iron-regulated genes in bacteria. *Microbiol Mol Biol Rev* 61:319–336
- Davies GR, Simmonds NJ, Stevens TR, Sheaff MT, Banatvala N, Laurenson IF, Blake DR, Rampton DS (1994) *Helicobacter pylori* stimulates antral mucosal reactive oxygen metabolite production in vivo. *Gut* 35:179–185
- Dorman CJ (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* 2:391–400
- Dunn BE, Cohen H, Blaser MJ (1997) *Helicobacter pylori*. *Clin Microbiol Rev* 10:720–741
- Dussurget O, Dumas E, Archambaud C, Chafsey I, Chambon C, Hébraud M, Cossart P (2005) *Listeria monocytogenes* ferritin protects against multiple stresses and is required for virulence. *FEMS Microbiol Lett* 250:253–261
- Eriksson S, Lucchini S, Thompson A, Rhen M, Hinton JC (2003) Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. *Mol Microbiol* 47:103–118
- Evans DJ Jr, Evans DG, Takemura T, Nakano H, Lampert HC, Graham DY, Granger DN, Kvietys PR (1995) Characterization of a *Helicobacter pylori* neutrophil-activating protein. *Infect Immun* 63:2213–2320
- Farr SB, Kogoma T (1991) Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*. *Microbiol Rev* 55:561–585
- Fiorini F, Stefanini S, Valenti P, Chiancone E, De Biase D (2008) Transcription of the *Listeria monocytogenes* *fri* gene is growth-phase dependent and is repressed directly by Fur, the ferric uptake regulator. *Gene* 410:113–121
- Frees D, Savijoki K, Varmanen P, Ingmer H (2007) Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Mol Microbiol* 63:1285–1289
- Frenkiel-Krispin D, Levin-Zaidman S, Shimoni E, Wolf SG, Wachtel EJ, Arad T, Finkel SE, Kolter R, Minsky A (2001) Regulated phase transitions of bacterial chromatin: a non-enzymatic pathway for generic DNA protection. *EMBO J* 20:1184–1191
- Frenkiel-Krispin D, Ben-Avraham I, Englander J, Shimoni E, Wolf SG, Minsky A (2004a) Nucleoid restructuring in stationary-state bacteria. *Mol Microbiol* 51:395–405
- Frenkiel-Krispin D, Sack R, Englander J, Shimoni E, Eisenstein M, Bullitt E, Horowitz-Scherer R, Hayes CS, Setlow P, Minsky A, Wolf SG (2004b) Structure of the DNA-SspC complex: implications for DNA packaging, protection, and repair in bacterial spores. *J Bacteriol* 186:3525–3530
- Fuangthong M, Helmann JD (2003) Recognition of DNA by three ferric uptake regulator (Fur) homologs in *Bacillus subtilis*. *J Bacteriol* 185:6348–6357
- Gangaidzo IT, Moyo VM, Mvundura E, Aggrey G, Murphree NL, Khumalo H, Saungweme T, Kasvosve I, Gomo ZA, Rouault T, Boelaert JR, Gordeuk VR (2001) Association of pulmonary tuberculosis with increased dietary iron. *J Infect Dis* 184:936–939
- Grainger DC, Goldberg MD, Lee DJ, Busby SJ (2008) Selective repression by Fis and H-NS at the *Escherichia coli* *dps* promoter. *Mol Microbiol* 68:1366–1377
- Grant RA, Filman DJ, Finkel SE, Kolter R, Hogle JM (1998) The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat Struct Biol* 5:294–303

- Grove A, Wilkinson SP (2005) Differential DNA binding and protection by dimeric and dodecameric forms of the ferritin homolog Dps from *Deinococcus radiodurans*. *J Mol Biol* 347:495–450
- Gupta S, Chatterji D (2003) Bimodal protection of DNA by *Mycobacterium smegmatis* Dps. *J Biol Chem* 278:5235–5241
- Gupta S, Pandit SB, Srinivasan N, Chatterji D (2002) Proteomics analysis of carbon-starved *Mycobacterium smegmatis*: induction of Dps-like protein. *Protein Eng* 15:503–511
- Halsey TA, Vazquez-Torres A, Gravidahl DJ, Fang FC, Libby SJ (2004) The ferritin-like Dps protein is required for *Salmonella enterica* serovar Typhimurium oxidative stress resistance and virulence. *Infect Immun* 72:1155–1158
- Hanna ES, Roque-Barreira MC, Bernardes ES, Panunto-Castelo A, Sousa MV, Almeida IC, Brocchi M (2007) Evidence for glycosylation on a DNA-binding protein of *Salmonella enterica*. *Microb Cell Fact* 6:11
- Hanna ES, Roque-Barreira MC, Mendes GM, Soares SG, Brocchi M (2008) Cloning, expression and purification of a glycosylated form of the DNA-binding protein Dps from *Salmonella enterica* Typhimurium. *Protein Exp Purif* 59:197–202
- Harrison PM, Arosio P (1996) The ferritins: molecular properties, iron storage function and cellular regulation. *Biochim Biophys Acta* 1275:161–203
- Hartford T, O'Brien S, Andrew PW, Jones D, Roberts IS (1993) Utilization of transferrin-bound iron by *Listeria monocytogenes*. *FEMS Microbiol Lett* 108:311–318
- Hatfield GW, Benham CJ (2002) DNA topology-mediated control of global gene expression in *Escherichia coli*. *Annu Rev Genet* 36:175–203
- Havukainen H, Haataja S, Kauko A, Pulliainen AT, Salminen A, Haikarainen T, Finne J, Papageorgiou AC (2008) Structural basis of the zinc- and terbium-mediated inhibition of ferroxidase activity in Dps ferritin-like proteins. *Protein Sci* 17:1513–1521
- Hébraud M, Guzzo J (2000) The main cold shock protein of *Listeria monocytogenes* belongs to the family of ferritin-like proteins. *FEMS Microbiol Lett* 190:29–34
- Henle ES, Linn S (1997) Formation, prevention, and repair of DNA damage by iron/hydrogen peroxide. *J Biol Chem* 272:19095–19098
- Herbig AF, Helmann JD (2001) Roles of metal ions and hydrogen peroxide in modulating the interaction of the *Bacillus subtilis* PerR peroxide regulon repressor with operator DNA. *Mol Microbiol* 41:849–859
- Higuchi M, Yamamoto Y, Kamio Y (2000) Molecular biology of oxygen tolerance in lactic acid bacteria: functions of NADH oxidases and Dpr in oxidative stress. *J Biosci Bioeng* 90:484–493
- Hurst JK, Barrette WC Jr (1989) Leukocytic oxygen activation and microbicidal oxidative toxins. *Crit Rev Biochem Mol Biol* 24:271–328
- Ilari A, Stefanini S, Chiancone E, Tsernoglou D (2000) The dodecameric ferritin from *Listeria innocua* contains a novel intersubunit iron-binding site. *Nat Struct Biol* 7:38–43
- Ilari A, Ceci P, Ferrari D, Rossi GL, Chiancone E (2002) Iron incorporation into *Escherichia coli* Dps gives rise to a ferritin-like microcrystalline core. *J Biol Chem* 277:37619–37623
- Ishikawa T, Mizunoe Y, Kawabata S, Takade A, Harada M, Wai SN, Yoshida S (2003) The iron-binding protein Dps confers hydrogen peroxide stress resistance to *Campylobacter jejuni*. *J Bacteriol* 185:1010–1017
- Kathariou S (2002) *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. *J Food Prot* 65:1811–1829
- Kauko A, Pulliainen AT, Haataja S, Meyer-Klaucke W, Finne J, Papageorgiou AC (2006) Iron incorporation in *Streptococcus suis* Dps-like peroxide resistance protein Dpr requires mobility in the ferroxidase center and leads to the formation of a ferrihydrite-like core. *J Mol Biol* 364:97–109
- Kim SG, Bhattacharyya G, Grove A, Lee YH (2006) Crystal structure of Dps-I, a functionally distinct Dps protein from *Deinococcus radiodurans*. *J Mol Biol* 361:105–111
- Kottakis F, Papadopoulos G, Pappa EV, Cordopatis P, Pentas S, Choli-Papadopoulou T (2008) *Helicobacter pylori* neutrophil-activating protein activates neutrophils by its C-terminal region

- even without dodecamer formation, which is a prerequisite for DNA protection-novel approaches against *Helicobacter pylori* inflammation. FEBS J 275:302–317
- Lee JW, Helmann JD (2007) Functional specialization within the Fur family of metalloregulators. *Biomaterials* 20:485–499
- Lee C, Lee SM, Mukhopadhyay P, Kim SJ, Lee SC, Ahn WS, Yu MH, Storz G, Ryu SE (2004) Redox regulation of OxyR requires specific disulfide bond formation involving a rapid kinetic reaction path. *Nat Struct Mol Biol* 11:1179–1185
- Levin-Zaidman S, Englander J, Shimoni E, Sharma AK, Minton KW, Minsky A (2003) Ringlike structure of the *Deinococcus radiodurans* genome: a key to radioresistance? *Science* 299:254–256
- Li X, Pal U, Ramamoorthi N, Liu X, Desrosiers DC, Eggers CH, Anderson JF, Radolf JD, Fikrig E (2007) The Lyme disease agent *Borrelia burgdorferi* requires BB0690, a Dps homologue, to persist within ticks. *Mol Microbiol* 63:694–710
- Liu S, Graham JE, Bigelow L, Morse PD, Wilkinson BJ (2002) Identification of *Listeria monocytogenes* genes expressed in response to growth at low temperature. *Appl Environ Microbiol* 68:1697–1705
- Liu X, Kim K, Leighton T, Theil EC (2006) Paired *Bacillus anthracis* Dps (mini-ferritin) have different reactivities with peroxide. *J Biol Chem* 281:27827–27835
- Lomovskaya OL, Kidwell JP, Matin A (1994) Characterization of the σ^{38} -dependent expression of a core *Escherichia coli* starvation gene, *pexB*. *J Bacteriol* 176:3928–3935
- Lu AL, Li X, Gu Y, Wright PM, Chang DY (2001) Repair of oxidative DNA damage: mechanisms and functions. *Cell Biochem Biophys* 35:141–170
- Luo Y, Han Z, Chin SM, Linn S (1994) Three chemically distinct types of oxidants formed by iron-mediated Fenton reactions in the presence of DNA. *Proc Natl Acad Sci USA* 91:12438–12442
- Martinez A, Kolter R (1997) Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J Bacteriol* 179:5188–5194
- Matin A (1991) The molecular basis of carbon-starvation-induced general resistance in *Escherichia coli*. *Mol Microbiol* 5:3–10
- Miller RA, Britigan BE (1997) Role of oxidants in microbial pathophysiology. *Clin Microbiol Rev* 10:1–18
- Mizuki I, Shimoyama T, Fukuda S, Liu Q, Nakaji S, Munakata A (2000) Association of gastric epithelial apoptosis with the ability of *Helicobacter pylori* to induce a neutrophil oxidative burst. *J Med. Microbiol* 49:521–524
- Mohamed W, Darji A, Domann E, Chiancone E, Chakraborty T (2006) The ferritin-like protein Frm is a target for the humoral immune response to *Listeria monocytogenes* and is required for efficient bacterial survival. *Mol Genet Genomics* 275:344–353
- Morikawa K, Ohniwa RL, Kim J, Maruyama A, Ohta T, Takeyasu K (2006) Bacterial nucleoid dynamics: oxidative stress response in *Staphylococcus aureus*. *Genes Cells* 11:409–423
- Morikawa K, Ohniwa RL, Kim J, Takeshita SL, Maruyama A, Inose Y, Takeyasu K, Ohta T (2007) Biochemical, molecular genetic, and structural analyses of the staphylococcal nucleoid. *Microsc Microanal* 13:30–35
- Nair S, Finkel SE (2004) Dps protects cells against multiple stresses during stationary phase. *J Bacteriol* 186:4192–4198
- Namavar F, Sparrius M, Veerman EC, Appelmelk BJ, Vandenbroucke-Grauls CM (1998) Neutrophil-activating protein mediates adhesion of *Helicobacter pylori* to sulfated carbohydrates on high-molecular-weight salivary mucin. *Infect Immun* 66:444–447
- Nicodème M, Perrin C, Hols P, Bracquart P, Gaillard JL (2004) Identification of an iron-binding protein of the Dps family expressed by *Streptococcus thermophilus*. *Curr Microbiol* 48:51–56
- Ohniwa RL, Morikawa K, Kim J, Ohta T, Ishihama A, Wada C, Takeyasu K (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. *EMBO J* 25:5591–5602
- Olczak AA, Olson JW, Maier RJ (2002) Oxidative-stress resistance mutants of *Helicobacter pylori*. *J Bacteriol* 184:3186–3193
- Olsen KN, Larsen MH, Gahan CG, Kallipolitis B, Wolf XA, Rea R, Hill C, Ingmer H (2005) The Dps-like protein Fri of *Listeria monocytogenes* promotes stress tolerance and intracellular multiplication in macrophage-like cells. *Microbiology* 151:925–933

- Papinutto E, Dundon WG, Pitulis N, Battistutta R, Montecucco C, Zanotti G (2002) Structure of two iron-binding proteins from *Bacillus anthracis*. *J Biol Chem* 277:15093–15098
- Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream MA, Rutherford KM, van Vliet AH, Whitehead S, Barrell BG (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 403:665–668
- Peña MMO, Bullerjahn GS (1995) The DpsA protein of *Synechococcus* sp. strain PCC7942 is a DNA-binding hemoprotein. *J Biol Chem* 270:22478–22482
- Phadtare S, Inouye M (2004) Genome-wide transcriptional analysis of the cold shock response in wild-type and cold-sensitive, quadruple-*csp*-deletion strains of *Escherichia coli*. *J Bacteriol* 186:7007–7014
- Phan-Thanh L, Gormon T (1995) Analysis of heat and cold shock proteins in *Listeria* by two-dimensional electrophoresis. *Electrophoresis* 16:444–450
- Pogliano K, Harry E, Losick R (1995) Visualization of the subcellular location of sporulation proteins in *Bacillus subtilis* using immunofluorescence microscopy. *Mol Microbiol* 18:459–470
- Polidoro M, de Biase D, Montagnini B, Guarrera L, Cavallo S, Valenti P, Stefanini S, Chiancone E (2002) The expression of the dodecameric ferritin in *Listeria* spp. is induced by iron limitation and stationary growth phase. *Gene* 296:121–128
- Pulliaainen AT, Haataja S, Kahkonen S, Finne J (2003) Molecular basis of H₂O₂ resistance mediated by Streptococcal Dpr. Demonstration of the functional involvement of the putative ferroxidase center by site-directed mutagenesis in *Streptococcus suis*. *J Biol Chem* 278:7996–8005
- Pulliaainen AT, Kauko A, Haataja S, Papageorgiou AC, Finne J (2005) Dps/Dpr ferritin-like protein: insights into the mechanism of iron incorporation and evidence for a central role in cellular iron homeostasis in *Streptococcus suis*. *Mol Microbiol* 57:1086–1100
- Ragkousi K, Cowan AE, Ross MA, Setlow P (2000) Analysis of nucleoid morphology during germination and outgrowth of spores of *Bacillus* species. *J Bacteriol* 182:5556–5562
- Ramsay B, Wiedenheft B, Allen M, Gauss GH, Lawrence CM, Young M, Douglas T (2006) Dps-like protein from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Inorg Biochem* 100:1061–1068
- Rea R, Hill C, Gahan CG (2005) *Listeria monocytogenes* PerR mutants display a small-colony phenotype, increased sensitivity to hydrogen peroxide, and significantly reduced murine virulence. *Appl Environ Microbiol* 71:8314–8322
- Ren B, Tibbelin G, Kajino T, Asami O, Ladenstein R (2003) The multi-layered structure of Dps with a novel di-nuclear ferroxidase center. *J Mol Biol* 329:467–477
- Roy S, Saraswathi R, Gupta S, Sekar K, Chatterji D, Vijayan M (2007) Role of N and C-terminal tails in DNA binding and assembly in Dps: structural studies of *Mycobacterium smegmatis* Dps deletion mutants. *J Mol Biol* 370:752–767
- Roy S, Saraswathi R, Chatterji D, Vijayan M (2008) Structural studies on the second *Mycobacterium smegmatis* Dps: invariant and variable features of structure, assembly and function. *J Mol Biol* 375:948–959
- Saumaa S, Tover A, Tark M, Tegova R, Kivisaar M (2007) Oxidative DNA damage defense systems in avoidance of stationary-phase mutagenesis in *Pseudomonas putida*. *J Bacteriol* 189:5504–5514
- Schnetz K (2008) Fine-tuned growth phase control of *dps*, encoding a DNA protection protein, by FIS and H-NS. *Mol Microbiol* 68:1345–1347
- Setlow B, Hand AR, Setlow P (1991) Synthesis of a *Bacillus subtilis* small, acid-soluble spore protein in *Escherichia coli* causes cell DNA to assume some characteristics of spore DNA. *J Bacteriol* 173:1642–1653
- Spurio R, Dürrenberger M, Falconi M, La Teana A, Pon CL, Gualerzi CO (1992) Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. *Mol Gen Genet* 231:201–211
- Stefanini S, Cavallo S, Montagnini B, Chiancone E (1999) Incorporation of iron by the unusual dodecameric ferritin from *Listeria innocua*. *Biochem J* 338:71–75

- Stephani K, Weichart D, Hengge R (2003) Dynamic control of Dps protein levels by ClpXP and ClpAP proteases in *Escherichia coli*. Mol Microbiol 49:1605–1614
- Stillman TJ, Upadhyay M, Norte VA, Sedelnikova SE, Carradus M, Tzokov S, Bullough PA, Shearman CA, Gasson MJ, Williams CH, Artymiuk PJ, Green J (2005) The crystal structures of *Lactococcus lactis* MG1363 Dps proteins reveal the presence of an N-terminal helix that is required for DNA binding. Mol Microbiol 57:1101–1112
- Storz G, Tartaglia LA, Farr SB, Ames BN (1990) Bacterial defenses against oxidative stress. Trends Genet 6:363–368
- Su M, Cavallo S, Stefanini S, Chiancone E, Chasteen ND (2005) The so-called *Listeria innocua* ferritin is a Dps protein. Iron incorporation, detoxification, and DNA protection properties. Biochemistry 44:5572–5578
- Sword CP (1966) Mechanisms of pathogenesis in *Listeria monocytogenes* infection. Influence of iron. J Bacteriol 92:536–542
- Tendeng C, Bertin PN (2003) H-NS in Gram-negative bacteria: a family of multifaceted proteins. Trends Microbiol 11:511–518
- Teneberg S, Miller-Podraza H, Lampert HC, Evans DJ Jr, Evans DG, Danielsson D, Karlsson KA (1997) Carbohydrate binding specificity of the neutrophil-activating protein of *Helicobacter pylori*. J Biol Chem 272:19067–19071
- Tikkanen K, Haataja S, François-Gerard C, Finne J (1995) Purification of a galactosyl-alpha 1-4-galactose-binding adhesin from the Gram-positive meningitis-associated bacterium *Streptococcus suis*. J Biol Chem 270:28874–28878
- Tonello F, Dundon WG, Satin B, Molinari M, Tognon G, Grandi G, Del Giudice G, Rappuoli R, Montecucco C (1999) The *Helicobacter pylori* neutrophil-activating protein is an iron-binding protein with dodecameric structure. Mol Microbiol 34:238–246
- Ueshima J, Shoji M, Ratnayake DB, Abe K, Yoshida S, Yamamoto K, Nakayama K (2003) Purification, gene cloning, gene expression, and mutants of Dps from the obligate anaerobe *Porphyromonas gingivalis*. Infect Immun 71:1170–1178
- Valdivia RH, Falkow S (1996) Bacterial genetics by flow cytometry: rapid isolation of *Salmonella typhimurium* acid-inducible promoters by differential fluorescence induction. Mol Microbiol 22:367–378
- van Vliet AH, Baillon ML, Penn CW, Ketley JM (1999) *Campylobacter jejuni* contains two fur homologs: characterization of iron-responsive regulation of peroxide stress defense genes by the PerR repressor. J Bacteriol 181:6371–6376
- Vazquez-Torres A, Fang FC (2001) *Salmonella* evasion of the NADPH phagocyte oxidase. Microbes Infect 3:1313–1320
- Velayudhan J, Castor M, Richardson A, Main-Hestor KL, Fang FC (2007) The role of ferritins in the physiology of *Salmonella enterica* sv. Typhimurium: a unique role for ferritin B in iron-sulphur cluster repair and virulence. Mol Microbiol 63:1495–1507
- Wang CA, Liu YC, Du SY, Lin CW, Fu HW (2008) *Helicobacter pylori* neutrophil-activating protein promotes myeloperoxidase release from human neutrophils. Biochem Biophys Res Commun 377:52–56
- Wen Y, Marcus EA, Matrubutham U, Gleeson MA, Scott DR, Sachs G (2003) Acid-adaptive genes of *Helicobacter pylori*. Infect Immun 71:5921–5939
- Wiedenheft B, Mosolf J, Willits D, Yeager M, Dryden KA, Young M, Douglas T (2005) An archaeal antioxidant: characterization of a Dps-like protein from *Sulfolobus solfataricus*. Proc Natl Acad Sci USA 102:10551–10556
- Wolf SG, Frenkiel D, Arad T, Finkel SE, Kolter R, Minsky A (1999) DNA protection by stress-induced biocrystallization. Nature 400:83–85
- Wooldridge KG, Williams PH (1993) Iron uptake mechanisms of pathogenic bacteria. FEMS Microbiol Rev 12:325–348
- Yamamoto Y, Higuchi M, Poole LB, Kamio Y (2000) Role of the *dpr* product in oxygen tolerance in *Streptococcus mutans*. J Bacteriol 182:3740–3747
- Yamamoto Y, Poole LB, Hantgan RR, Kamio Y (2002) An iron-binding protein, Dpr, from *Streptococcus mutans* prevents iron-dependent hydroxyl radical formation in vitro. J Bacteriol 184:2931–2939

- Yamamoto Y, Fukui K, Koujin N, Ohya H, Kimura K, Kamio Y (2004) Regulation of the intracellular free iron pool by Dpr provides oxygen tolerance to *Streptococcus mutans*. J Bacteriol 186:5997–6002
- Yoo AY, Kim SW, Yu JE, Kim YH, Cha J, Oh JI, Eo SK, Lee JH, Kang HY (2007) Requirement of Fur for the full induction of Dps expression in *Salmonella enterica* serovar Typhimurium. J Microbiol Biotechnol 17:1452–1459
- Yoshida N, Granger DN, Evans DJ Jr, Evans DG, Graham DY, Anderson DC, Wolf RE, Kvietys PR (1993) Mechanisms involved in *Helicobacter pylori*-induced inflammation. Gastroenterology 105:1431–1440
- Zanotti G, Papinutto E, Dundon W, Battistutta R, Seveso M, Giudice G, Rappuoli R, Montecucco C (2002) Structure of the neutrophil-activating protein from *Helicobacter pylori*. J Mol Biol 323:125–130
- Zhao G, Ceci P, Ilari A, Giangiacomo L, Laue TM, Chiancone E, Chasteen ND (2002) Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells. A ferritin-like DNA-binding protein of *Escherichia coli*. J Biol Chem 277:27689–27696
- Zheng M, Aslund F, Storz G (1998) Activation of the OxyR transcription factor by reversible disulfide bond formation. Science 279:1718–1721

Part II
Chromatin Organization in Archaea
and Eukaryotes

Chapter 10

Archaeal Chromatin Organization

Stephen D. Bell and Malcolm F. White

Abstract In the light microscope, archaea resemble bacteria, in that they are small, generally single-celled organisms devoid of overt subcellular organization. Their genomes (which range in size from 0.5 to ~6 Mb) are found in condensed nucleoid structures, rather than within nuclei and so the archaea can be broadly classified as “prokaryotes”. As in bacteria, a complex variety of proteins appear to play roles in compacting archaeal nucleoid structures. However, despite the organisational similarity between bacterial and archaeal subcellular features, archaeal nucleoid-associated proteins have intriguing parallels with the proteins that shape eukaryotic chromatin. These similarities manifest themselves both at the physical level, in the form of the structural orthology of some eukaryotic and archaeal histone proteins, and at the conceptual level, in the role of covalent modifications in modulating the DNA binding mode of archaeal chromatin proteins.

It has been recognised for some time that two principal phyla of archaea exist – the Crenarchaea and the Euryarchaea. Recently, it has been proposed that what was previously believed to be a divergent branch of the Crenarchaea may in fact represent a novel third archaeal phylum, the Thaumarchaea (Brochier-Armanet et al. 2008). In addition the genome of a candidate member of a fourth phylum, the Korarchaeota, has recently been sequenced (Elkins et al. 2008). Interestingly, some archaeal proteins show characteristic distributions across these four phyla. However, in contrast to the ubiquity of histones in eukaryotes, there are no chromatin proteins that are universally conserved in the archaeal domain of life. In the following chapter we will describe the principal archaeal chromatin proteins in a variety of model species.

S.D. Bell (✉)

Sir William Dunn School of Pathology, Oxford University,
South Parks Rd, Oxford, OX1 3RE, UK
e-mail: stephen.bell@path.ox.ac.uk

M.F. White

Biomolecular Sciences Building, University of St Andrews,
North Haugh, St. Andrews, Fife, KY16 9ST, UK
e-mail: mfw2@st-andrews.ac.uk

10.1 Sul7

Members of the Sul7 family of chromatin proteins are found in the highly studied crenarchaeal *Sulfolobus* species where they comprise about 5% of the total cellular protein and also in the related *Metallosphaera sedula*. While *M. sedula* has a single *sul7* gene, two *sul7* genes are found in *S. tokodaii* and *S. acidocaldarius*. These genes encode identical proteins in *S. tokodaii* but there are five amino acid coding differences between the two *S. acidocaldarius* copies. The family has undergone a limited expansion in *S. solfataricus*, with three genes, two of which encode identical proteins and the third a species that differs in a single amino acid. None of the amino acid substitutions are in any of the residues important in contacting DNA; whether the different isoforms have distinct properties and roles in the cell remains unclear. The Sul7 proteins are small (~7 kDa), monomeric and highly basic. A number of structures of Sul7 family proteins have been determined, both of the free protein and in the context of a protein-DNA complex (Edmondson and Shriver 2001; Gao et al. 1998; Krueger et al. 1999). The protein resembles an SH3-like organisation with a β -barrel fold. Residues within the β -barrel make DNA contacts – binding to bases via the minor groove (Fig. 10.1).

Strikingly, residues V26 and M29 intercalate between bases in the DNA, thereby introducing a sharp kink in the double helix that results in a bend of about 66° (Gao et al. 1998). This under-winding and bending provides the basis for the known ability of Sul7 proteins to constrain negative supercoils in vitro (Napoli et al. 2002).

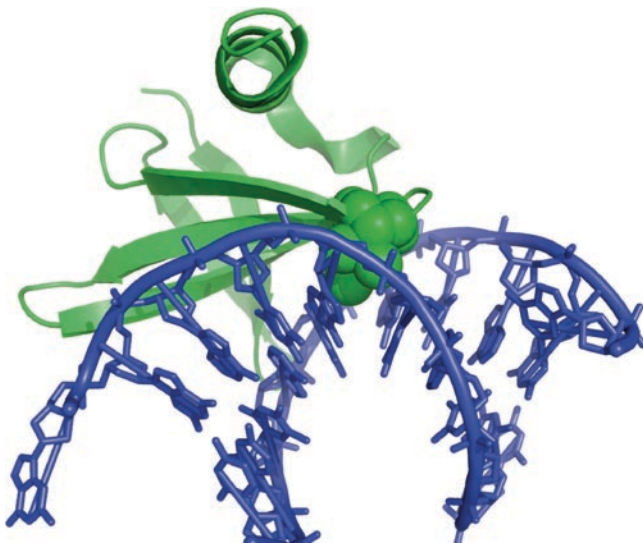


Fig. 10.1 The structure of Sul7 bound to DNA. The intercalating residues V26 and M29 are shown as spheres. The figure was generated with Pymol using PDB coordinates 1AZQ

The abundance of these proteins in *Sulfolobus* species, coupled with their non-specific DNA binding activity, suggests that they may play a major role in nucleoid compaction or organisation. However, while sub-cellular fractionation data support a general association of Sul7 proteins with the nucleoid (Napoli et al. 2004), little is known about their localisation along the genome. One intriguing feature of the native Sul7 proteins is that they are methylated on lysine residues (Edmondson and Shriver 2001). The consequence of this remains poorly understood; methylated and non-methylated forms of the protein bind DNA with the same affinity (Edmondson and Shriver 2001). Methylation has been observed to increase following heat shock and so it is possible that it may contribute to the thermostability of the protein (Edmondson and Shriver 2001). In light of the methylation of the protein, it is particularly enticing that the fold of Sul7 resembles that of the methyl-lysine binding chromodomain found in eukaryotic histones (Jacobs and Khorasanizadeh 2002; Nielsen et al. 2002). However, while it is tempting to speculate that some form of regulatory oligomerisation may be mediated by differential methylation of Sul7, there are currently no data to support this hypothesis.

In addition to its likely role in chromatin compaction, Sul7 also may play two additional roles in stabilising DNA. The first is in helping to maintain DNA duplex integrity at elevated temperature. DNA complexed with Sul7 shows an increase in melting temperature of up to 40°C, leading to speculation that Sul7 family members may coat or sheath the DNA, thus helping to hold base-pairs together (McAfee et al. 1996). A second, and much more surprising activity attributed to Sul7 lies in its ability to protect DNA from UV-induced damage. More specifically, it was shown that Sul7 has the capacity to reverse thymidine dimers. This ability is manifested in the form of a photoreactivation type reaction in which irradiation of Sul7 leads to oxidation of a conserved tryptophan (W24) located at the DNA binding interface. However, if the tryptophan is juxtaposed to a thymidine dimer then an electron transfer occurs leading to oxidation and thus repair of the thymidine dimer (Tashiro et al. 2006). Although this process is much less efficient than photoreactivation mediated by photolyases, the very high concentration of Sul7 in vivo may ensure that Sul7-mediated repair is physiologically relevant.

10.2 Cren7: A Structural Homolog of Sul7

Recent work has revealed that Sul7 may be a representative of an extended family of sequence divergent but nevertheless structurally related proteins. Huang and colleagues purified an abundant (~1% total cell protein) small protein from *Sulfolobus shibate*. The protein, termed Cren7, was found to be conserved in almost all crenarchaea, had non-specific dsDNA binding activity and could constrain negative supercoils in vitro (Guo et al. 2008). Like Sul7d, it displayed monomethylation at multiple sites. However, as with Sul7, no functional significance could be attributed to the methylation. The parallels in biochemical behaviour between Cren7 and Sul7 were underscored by the determination of the structure of Cren7. Remarkably,

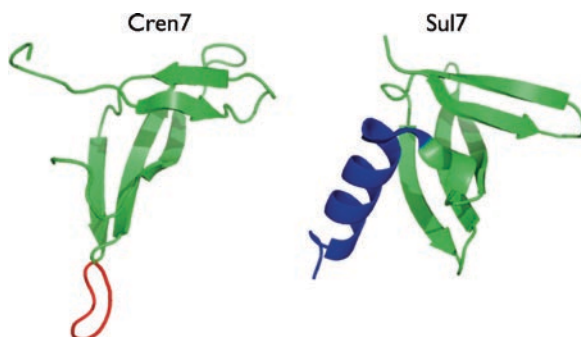


Fig. 10.2 Comparison of Cren7 and Sul7. The extended loop in Cren7 is shown in red and the extra C-terminal α -helix of Sul7 is in blue. The figure was generated with Pymol using PDB coordinates 2JTM (Cren7) and 1SSO (Sul7)

despite the complete lack of significant sequence similarity between the proteins, they shared the same basic fold by having a β -barrel based SH3-like structure (Guo et al. 2008). As can be seen in Fig. 10.2, the principal differences lie in the presence of an extended loop in Cren7 (shown in red) and an extra C-terminal α -helix in Sul7 (in blue). NMR studies suggested that DNA binding by Cren7 involves an equivalent surface to that employed by Sul7, significantly the important hydrophobic intercalating residues (V26 and M29) in Sul7 have conservative replacements (L28 and V36) in Cren7.

10.3 CC1

Examination of the genome sequences of organisms from the order *Thermoproteales* revealed a surprising absence of clear homologs of single-stranded DNA binding proteins (SSBs). Although of diverse sequence and composition across the three domains of life, SSBs share a structural signature in the form of an OB fold. The lack of this otherwise universal fold in *Thermoproteales* prompted a biochemical search for candidate ssDNA binding activities in organisms of this order. This led to the identification of CC1, a small basic protein of primarily beta-sheet organisation (Luo et al. 2007). The protein has the capacity to bind with high cooperativity and with nearly identical affinities to either ssDNA or dsDNA (Luo et al. 2007) and a binding site size of approximately 6 bp per monomer (Hardy and Martin 2008). The ability to bind tightly to ssDNA differentiates CC1 from Sso7d, which binds ssDNA very poorly (Hardy and Martin 2008). While CC1 is restricted to the *Thermoproteales* and *Aeropyrum pernix*, the beta-sheet rich structure is reminiscent of both Sul7 and Cren7, suggesting that despite low amino acid sequence similarity, there may turn out to be structural homologies between CC1 and these proteins. We have therefore grouped these proteins together in Fig. 10.7.

10.4 MC1

MC1 is an abundant DNA binding protein in species of methanogenic euryarchaea of the order *Methanosarcinales*. It is additionally found in some halophilic euryarchaea such as *Haloquadratum*, *Halobacterium*, *Haloarcula* and *Natromonas*. While single genes for MC1 are found in the halophiles, *Methanosarcina* species have two MC1 genes, encoding proteins that are about 90% identical. As with many chromatin-associated proteins, MC1 proteins are small (87–94 amino acids) and highly basic. They bind DNA non-cooperatively and each monomer covers about 10–11 bp. Binding introduces a sharp bend of about 120° in DNA (Le Cam et al. 1999). In common with some bacterial chromatin proteins and eukaryotic HMG proteins, MC1 shows some preference for binding to aberrant DNA structures, such as mini-circles and four-way junctions (Teyssier et al. 1996; Toulme et al. 1995). The structure of *Methanosarcina thermophila* MC1 revealed that it had a novel fold with one α helix and five β -strands arranged in two antiparallel sheets that fold to form a β -barrel like structure (Paquet et al. 2004) (Fig. 10.3). The β -barrel shares topological similarities to that of the Sul7/Cren7 structures, although the loops connecting the β -strands vary significantly in size.

The precise DNA-binding mode employed by MC1 remains unclear, however, it is possible that it may resemble that of Sul7. Thus, the β -barrel fold employed by a number of archaeal chromatin proteins may be reflective of an ancestral DNA-organising domain that has been embellished in a number of archaeal lineages. Interestingly, *Methanosarcina thermophila* encodes a SET-domain protein. SET domains have methyltransferase activity and play pivotal roles in the differential methylation of eukaryotic chromatin proteins. Biochemical studies revealed that the *Methanosarcina* SET domain protein could methylate MC1, however, the physiological relevance of this in vitro observation remains unclear (Manzur and Zhou 2005).



Fig. 10.3 The structure of MC1. Comparison with Fig. 10.2 reveals the similar β -barrel folds of MC1, Cren7 and Sul7. The figure was generated with Pymol using PDB coordinates 1T23

10.5 Sac10a

Sac10a was isolated from *S. acidocaldarius* extracts as an abundant DNA binding protein that has the capacity to introduce supercoiling and bring together two duplexes into interwound structures. It has been estimated to be highly abundant, reaching levels similar to those of Sul7. DNA binding studies have revealed a modest sequence preference for A/T rich DNA (Edmondson et al. 2004). Given this modest degree of sequence preference of the protein, it is currently unclear whether the Sul10a protein is a true general chromatin protein, or a transcription factor associated with specific genetic loci. The protein is found in a range of both cren- and euryarchaea. Essentially nothing is known about the physiological role of this protein, beyond the observation that its abundance appears to reduce as *Sulfolobus* cultures enter stationary phase (Edmondson et al. 2004). Biochemical and structural studies have revealed the protein to be a homodimer. The protein homodimerises via a long anti-parallel coiled coil generating an extended structure (Kahsai et al. 2005). The linear extremities of the dimer are formed by winged-helix DNA binding domains, one contributed by each monomer (Fig. 10.4).

10.6 Alba

Identified at the same time as Sac10a, Alba (Sac10b or Sso10b) has been the subject of a number of biochemical and structural studies. Alba is found in both euryarchaeal and crenarchaeal kingdoms as well as in a variety of eukaryotes, ranging from Trypanosomes to Man (Bell et al. 2002). Several archaeal species, including members of the *Sulfolobus* genus, encode two Alba family members, Alba1 and Alba2 (Jelinska et al. 2005). In *Sulfolobus*, Alba1 comprises about 4% of total cell protein, while Alba2 is about 20-fold less abundant (Jelinska et al. 2005). The proteins function as dimers and structures of homodimers of Alba1 from a number of archaea have been determined. In addition the structure of a heterodimer of *Sulfolobus* Alba1–Alba2 has also been solved (Jelinska et al. 2005; Wardleworth et al. 2002). Interestingly the preferred higher order form of Alba2 appears to be in a heterodimer with Alba1. The structures reveal that Alba forms a twofold symmetric structure with a central alpha-helical rich core from which two



Fig. 10.4 Structure of the Sac10a homo-dimer. The figure was generated with Pymol using PDB coordinates 1R7J

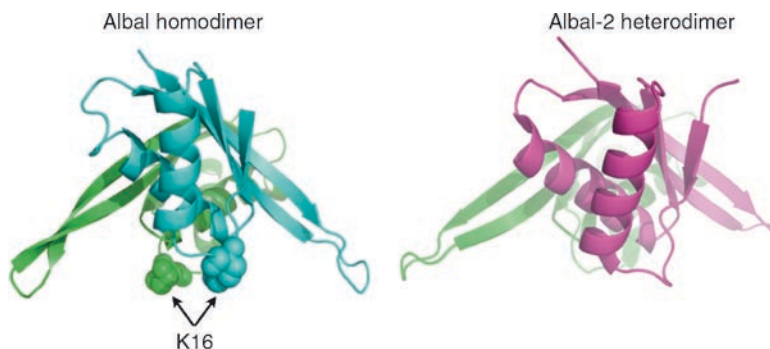


Fig. 10.5 Comparison of Alba1 homodimer and Alba1-2 heterodimer (*Alba1* in green, *Alba2* in magenta). The site of in vivo acetylation (K16) in Alba1 is shown as spheres. The figure was generated with Pymol using PDB coordinates 1HOX (Alba1 homodimer) and 2BKY (heterodimer)

β -hairpin wings protrude (Fig. 10.5). The fold of the N-terminal region of Alba resembles the DNA binding fold of DNaseI and also bears similarity to the RNA binding fold of translation factor IF3.

Alba appears to be a reasonably promiscuous nucleic acid binding protein with a preference for binding double stranded nucleic acids. Both DNA and RNA can be bound in vitro and in vivo (Guo et al. 2003; Marsh et al. 2005). Interaction of Alba1 with dsDNA constrains negative supercoils and electron microscopy studies have revealed that it coats DNA in a filamentous structure (Jelinska et al. 2005; Lurz et al. 1986). The nature of these filaments varies with changing Alba:DNA ratios. At 6 bp/Alba1 dimer, extended filaments are observed in which all the DNA is coated. At lower ratios (12 bp/dimer) tangled structures are observed in which loops of uncoated DNA are seen to extrude from stem-like structures in which two duplexes appear to be interwound by Alba-DNA filaments. The behaviour of Alba1:Alba2 heterodimers was similar to Alba1 homodimers at 12 bp/dimer. However, at higher protein concentrations the heterodimers resulted in the formation of novel, branched structures in which protein coated loops extruded out from a central condensed mass of protein and DNA (Jelinska et al. 2005). The ability of Alba to form these filament structures may explain its ability to raise the melting temperature of DNA significantly (Richard et al. 2004). This suggests, that in addition to organising chromatin, Alba may also act as a thermo-protectant for both DNA and RNA in the cell.

The crystallographic studies reveal a possible basis for the distinct behaviours of the homo- and heterodimers. Examination of the crystal lattices of Alba1 homodimers reveals that Alba1 has the propensity to form filament structures via end-to-end association of homodimers involving a strikingly well conserved interface. Significantly, this site is poorly conserved in the Alba2 protein, suggesting it may act as a filament breaker (Jelinska et al. 2005).

Native Alba1 from *S. solfataricus* was found to have lower DNA binding affinity than recombinant material purified from *Escherichia coli*. Mass spectrometry revealed that the native material possessed two sites of acetylation, one at the

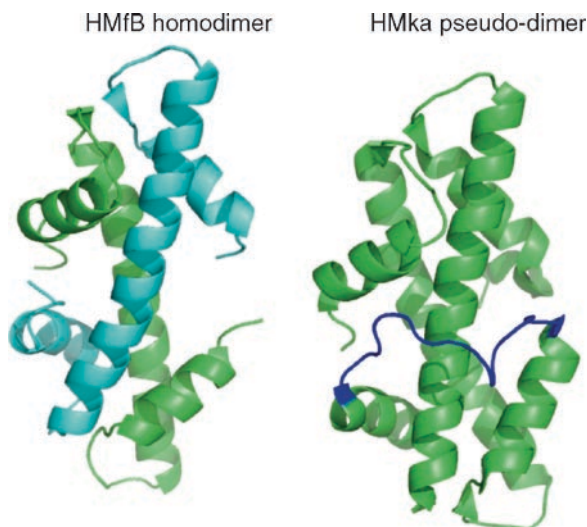


Fig. 10.6 Structure of a homodimer of *M. fervidus* HmfB and the pseudo-dimeric histone from *M. kandleri* (HmkA). The peptide linking the two histone folds is highlighted in blue. The figure was generated with Pymol using PDB coordinates 1A7W (HMfB) and 1F1E (HMkA)

N-terminus and the other at lysine 16 (Bell et al. 2002). Furthermore, mutagenesis of the K16 in the recombinant protein to either glutamate or alanine resulted in lowered DNA binding affinity. Thus it appears that acetylation of Alba may modulate its DNA binding affinity. The structure of Alba1 reveals that K16 is positioned in the alpha-helical bundle of Alba, in a position where it is likely to make direct contacts with DNA; removal of the positive charge of the lysine side chain upon acetylation would therefore be predicted to interfere with DNA binding.

It was particularly notable, therefore, that native Alba1 was found to co-immunopurify with *Sulfolobus* Sir2 – a homolog of the eukaryotic NAD-dependent Sir2 protein deacetylases. In eukaryotes, the Sir2 protein plays a number of significant roles in regulating the activity of proteins by reversible acetylation. Importantly, physiological substrates for eukaryotic Sir2 proteins include the histone proteins. Treatment of native Alba1 with Sir2 enhanced its DNA binding affinity and this could be detected by the enhanced repressive potential of Sir2-treated native Alba1 in *in vitro* transcription assays (Bell et al. 2002). Studies of bacterial Sir2 homologs had revealed that they play a key role in the regulation of acetyl coA synthetase (ACS) in partnership with an acetyltransferase, Pat (Starai et al. 2002; Starai and Escalante-Semerena 2004). Interestingly in both Alba1 and bacterial ACS, the acetylated lysine is preceded by a glycine. Furthermore, archaea encoded a short protein homologous to the acetyltransferase domain of Pat. Experiments with the *Sulfolobus* Pat homolog (ssPat), revealed that this protein was able to acetylate Alba1 and did so specifically on K16. Furthermore, this led to a reduction in the DNA binding affinity of Alba1 *in vitro* (Marsh et al. 2005). Thus, it appears likely that the Pat-Sir2 partnership, that in bacteria is involved in modulating the key

metabolic enzyme, acetyl CoA synthetase, has been co-opted in *Sulfolobus* to form a primitive form of chromatin regulation machinery. It is tempting to speculate that this may allow coupling of the global control of chromatin organisation (and thus DNA replication and gene transcription) to changes in the energy flux and metabolic status of cells.

10.7 Archaeal Histones

In 1990, studies of the DNA binding proteins of *Methanothermobacter thermautotrophicus* revealed that this species encoded bona fide orthologs of eukaryotic histone proteins (Sandman et al. 1990). The principal unit of DNA compaction in eukaryotes is the histone octamer. This contains two copies each of histone H2A, H2B, H3 and H4. The assembly of the octamer and its interaction with DNA to form a nucleosome core particle is an ordered and chaperoned event in eukaryotes. First a basic tetramer of 2 copies of each of H3 and H4 forms a tetramer, this then has a one copy of H2A and H2B added either side of the H3/H4 tetramer to form the histone octamer. The archaeal histones most closely resemble eukaryotic H3 and H4 (Reeve et al. 2004). All four eukaryotic histones have a common core fold, the “histone fold” containing three alpha helices that are separated by two short β -loops. In eukaryotes, tails are found both N- and C-terminal of the core histone fold. The tail sequences show considerable conservation within each histone type and contain a great many sites of post-translational modification. There is currently no evidence that any archaeal histones are subject to any post-translational modifications. Furthermore, most archaeal histones lack significant tails, at most they possess a few extra residues, beyond the core histone fold. A few archaeal histones do have short C-terminal tails (typically of less than 25 residues). These may play regulatory roles as deletion of one such tail (in *Methanocaldococcus jannaschii* MJ1647) appears to stimulate DNA binding by this protein (Li et al. 2000). While the precise physiological role of the histones remains poorly understood, a body of data has been amassed on the biochemistry of the proteins. In general, histone fold-containing proteins are insoluble as monomers but stable as dimers. Many archaea possess multiple histone homologs and these have been shown to be capable of both homo-dimeric and hetero-dimeric interactions. In two archaea, *Methanopyrus kandleri* and *Halobacterium salinarum*, unusual histones are observed which contain two histone folds within one polypeptide (Fahrner et al. 2001); (Sandman and Reeve 2006). Thus, these proteins can be viewed as obligate pseudo-heterodimers.

Many archaea encode multiple histone genes and in *Thermococcus zilligii*, *Methanococcus voltae* and *M. fervidus*, differential expression of distinct histone homologs during culture growth was observed leading to speculation that the differential abundance of the homo and heterodimeric forms may modulate the nature of nucleoid organisation in different growth conditions (Dinger et al. 2000; Sandman et al. 1994).

When binding to DNA, archaeal histones form a tetramer or dimer of dimers analogous to the eukaryotic (H3/H4)₂ tetrasome (Reeve et al. 2004). Mutations in

the dimer–dimer interface impair the ability of the histones to bind DNA. The histone tetramer binds approximately 90 bp of dsDNA, wrapping it around the surface of the histone tetramer. Not surprisingly, *in vitro* substrate selection studies revealed that archaeal histones, like their eukaryotic counterparts, show elevated binding to intrinsically curved DNA sequences (Bailey et al. 2002).

10.8 Phylogenetic Distribution of Archaeal Chromatin Proteins

The chromatin protein distribution shown in Fig. 10.7 emphasises that there is no “universal” archaeal chromatin protein and, furthermore, suggests that displacement of one chromatin protein by another appears to have occurred several times in

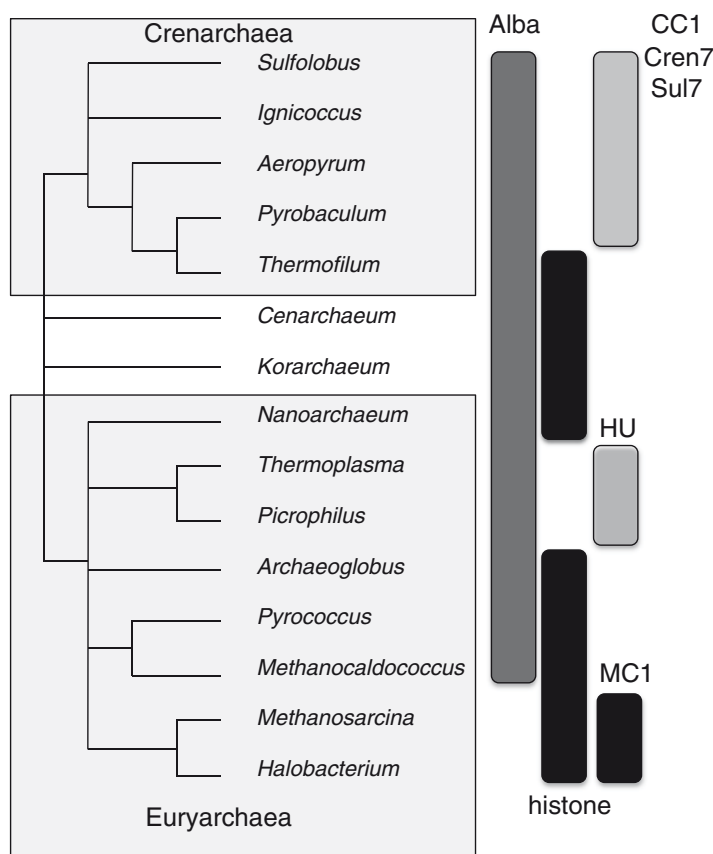


Fig. 10.7 Phylogenetic distribution of the major archaeal chromatin proteins. Representative archaeal species are indicated. *Cenarcheum* is a representative of the recently proposed Thaumarchaeal phylum

the course of evolution. The clearest example is in the *Thermoplasmatales*, where the archaeal histone has been lost, possibly due to lateral gene transfer of a bacterial gene encoding the chromatin protein HU. Similarly, it is tempting to postulate that Alba has been displaced by MC1 in the *Methanosarcinas* and halophiles. It is not possible to state definitively whether the Cren7-like family has displaced histones in most crenarchaea, or conversely that histone genes have been laterally transferred into a few crenarchaea such as *Thermophilum pendens*, resulting in loss of Cren7. What is clear is that all sequenced archaeal genomes encode at least two different chromatin proteins that presumably collaborate to organise and compact the genome (eg. the Alba and Cren7 families in *Sulfolobales*; Alba and HU in *Thermoplasmatales*; Histones and Alba in *Pyrococcales* etc.). A similar situation exists in the bacteria, and as in the bacteria no single archaeal chromatin protein may be essential for cell viability, as deletion of either Alba or histone genes has been demonstrated in *Methanococcus voltae* (Heinicke et al. 2004). Thus, archaeal chromatin presents a curious blend of eukaryotic and bacterial features. With the continuing improvements in archaeal genetic systems it is anticipated that there will be significant and rapid progress in elucidating the precise roles of archaeal chromatin proteins in shaping the organisation of the nucleoid in these fascinating organisms.

References

- Bailey KA, Marc F, Sandman K, Reeve JN (2002) Both DNA and histone fold sequences contribute to archaeal nucleosome stability. *J Biol Chem* 277:9293–9301
- Bell SD, Botting CH, Wardleworth BN, Jackson SP, White MF (2002) The interaction of Alba, a conserved archaeal, chromatin protein, with Sir2 and its regulation by acetylation. *Science* 296:148–151
- Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* 6:245–252
- Dinger ME, Baillie GJ, Musgrave DR (2000) Growth phase-dependent expression and degradation of histones in the thermophilic archaeon *Thermococcus zilligii*. *Mol Microbiol* 36:876–885
- Edmondson SP, Shriver JW (2001) DNA-binding proteins Sac7d and Sso7d from *Sulfolobus*. In: *Hyperthermophilic enzymes*, Pt C, vol 334, pp 129–145
- Edmondson SP, Kahsai MA, Gupta R, Shriver JW (2004) Characterization of Sac10a, a hyperthermophile DNA-binding protein from *Sulfolobus acidocaldarius*. *Biochemistry* 43:13026–13036
- Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L, Hedlund BP, Brochier-Armanet C, Kunin V, Anderson I, Lapidus A, Goltsman E, Barry K, Koonin EV, Hugenholtz P, Kyrpides N, Wanner G, Richardson P, Keller M, Stetter KO (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc Natl Acad Sci USA* 105:8102–8107
- Fahrner RL, Cascio D, Lake JA, Slesarev A (2001) An ancestral nuclear protein assembly: crystal structure of the *Methanopyrus kandleri* histone. *Protein Sci* 10:2002–2007
- Gao YG, Su SY, Robinson H, Padmanabhan S, Lim L, McCrary BS, Edmondson SP, Shriver JW, Wang AHJ (1998) The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. *Nat Struct Biol* 5:782–786
- Guo R, Xue H, Huang L (2003) Ssh10b, a conserved thermophilic archaeal protein, binds RNA in vivo. *Mol Microbiol* 50:1605–1615

- Guo L, Feng YG, Zhang ZF, Yao HW, Luo YM, Wang JF, Huang L (2008) Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. *Nucleic Acids Res* 36:1129–1137
- Hardy CD, Martin PK (2008) Biochemical characterization of DNA-binding proteins from *Pyrobaculum aerophilum* and *Aeropyrum pernix*. *Extremophiles* 12:235–246
- Heinicke I, Muller J, Pittelkow M, Klein A (2004) Mutational analysis of genes encoding chromatin proteins in the archaeon *Methanococcus voltae* indicates their involvement in the regulation of gene expression. *Mol Genet Genom* 272:76–87
- Jacobs SA, Khorasanizadeh S (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 295:2080–2083
- Jelinska C, Conroy MJ, Craven CJ, Hounslow AM, Bullough PA, Waltho JP, Taylor GL, White MF (2005) Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. *Structure* 13:963–971
- Kahsai MA, Vogler B, Clark AT, Edmondson SP, Shriver JW (2005) Solution structure, stability, and flexibility of Sso10a: A hyperthermophile coiled-coil DNA-binding protein. *Biochemistry* 44:2822–2832
- Krueger JK, McCrary BS, Wang AHJ, Shriver JW, Trewhella J, Edmondson SP (1999) The solution structure of the Sac7d/DNA complex: a small-angle X-ray scattering study. *Biochemistry* 38:10247–10255
- Le Cam E, Culard F, Larquet E, Delain E, Cognet JAH (1999) DNA bending induced by the archaeobacterial histone-like protein MC1. *J Mol Biol* 285:1011–1021
- Li WT, Sandman K, Pereira SL, Reeve JN (2000) MJ1647, an open reading frame in the genome of the hyperthermophile *Methanococcus jannaschii*, encodes a very thermostable archaeal histone with a C-terminal extension. *Extremophiles* 4:43–51
- Luo X, Schwarz-Linek U, Botting CH, Hensel R, Siebers B, White MF (2007) CC1, a novel crenarchaeal DNA binding protein. *J Bacteriol* 189:403–409
- Lurz R, Grote M, Dijk J, Reinhardt R, Dobrinski B (1986) Electron-microscopic study of DNA complexes with proteins from the archaeobacterium *Sulfolobus acidocaldarius*. *EMBO J* 5:3715–3721
- Manzur KL, Zhou MM (2005) An archaeal SET domain protein exhibits distinct lysine methyltransferase activity towards DNA-associated protein MC1- α . *FEBS Lett* 579:3859–3865
- Marsh VL, Peak-Chew SY, Bell SD (2005) Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. *J Biol Chem* 280:21122–21128
- McAfee JG, Edmondson SP, Zegar I, Shriver JW (1996) Equilibrium DNA binding of Sac7d protein from the hyperthermophile *Sulfolobus acidocaldarius*: fluorescence and circular dichroism studies. *Biochemistry* 35:4034–4045
- Napoli A, Zivanovic Y, Bocs C, Buhler C, Rossi M, Forterre P, Ciaramella M (2002) DNA bending, compaction and negative supercoiling by the architectural protein Sso7d of *Sulfolobus solfataricus*. *Nucleic Acids Res* 30:2656–2662
- Napoli A, Valenti A, Salerno V, Nadal M, Garnier F, Rossi M, Ciaramella M (2004) Reverse gyrase recruitment to DNA after UV light irradiation in *Sulfolobus solfataricus*. *J Biol Chem* 279:33192–33198
- Nielsen PR, Nietlispach D, Mott HR, Callaghan J, Bannister A, Kouzarides T, Murzin AG, Murzina NV, Laue ED (2002) Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* 416:103–107
- Paquet F, Culard F, Barbault F, Maurizot JC, Lancelot G (2004) NMR solution structure of the archaeobacterial chromosomal protein MC1 reveals a new protein fold. *Biochemistry* 43:14971–14978
- Reeve JN, Bailey KA, Li WT, Marc F, Sandman K, Soares DJ (2004) Archaeal histones: structures, stability and DNA binding. *Biochem Soc Trans* 32:227–230
- Richard DJ, Bell SD, White MF (2004) Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase. *Nucleic Acids Res* 32:1065–1074
- Sandman K, Reeve JN (2006) Archaeal histones and the origin of the histone fold. *Curr Opin Microbiol* 9:520–525

- Sandman K, Krzycki JA, Dobrinski B, Lurz R, Reeve JN (1990) Hmf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus-Fervidus*, is most closely related to histones. *Proc Natl Acad Sci USA* 87:5788–5791
- Sandman K, Grayling RA, Dobrinski B, Lurz R, Reeve JN (1994) Growth-phase-dependent synthesis of histones in the archaeon *Methanothermus-Fervidus*. *Proc Natl Acad Sci USA* 91:12624–12628
- Starai VJ, Escalante-Semerena JC (2004) Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in *Salmonella enterica*. *J Mol Biol* 340:1005–1012
- Starai VJ, Celic I, Cole RN, Boeke JD, Escalante-Semerena JC (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science* 298:2390–2392
- Tashiro R, Wang AHJ, Sugiyama H (2006) Photoreactivation of DNA by an archaeal nucleoprotein Sso7d. *Proc Natl Acad Sci USA* 103:16655–16659
- Teyssier C, Toulme F, Touzel JP, Maurizot JC, Culard F (1996) Preferential binding of the archaeobacterial histone-like MC1 protein to negatively supercoiled DNA minicircles. *Biochemistry* 35:7954–7958
- Toulme F, Lecam E, Teyssier C, Delain E, Sautiere P, Maurizot JC, Culard F (1995) Conformational changes of DNA minicircles upon the binding of the archaeobacterial histone-like protein MC1. *J Biol Chem* 270:6286–6291
- Wardleworth BN, Russell RJM, Bell SD, Taylor GL, White MF (2002) Structure of Alba: an archaeal chromatin protein modulated by acetylation. *EMBO J* 21:4654–4662

Chapter 11

The Topology and Organization of Eukaryotic Chromatin

Andrew Travers and Georgi Muskhelishvili

Abstract In all organisms DNA is maintained in a highly compacted state complexed with abundant basic proteins. The extent of this compaction can range from ~1,000-fold in the bacterial nucleoid to ~10,000-fold in eukaryotic metaphase chromosomes. In addition to the necessity for compaction the genetic specification function of DNA also requires that the appropriate encoded information be accessible for transcription. The dual requirements of compaction and selective accessibility imply that the complex of DNA and abundant basic proteins, defined here generally as chromatin, must possess a high degree of structural organisation and that the regulation of transcription at the level of the gene may involve substantial structural transitions. In this article we summarise the current understanding of the organisation of eukaryotic chromatin and discuss the extent to which the organisational principles are also apparent in prokaryotic chromatin.

Keywords Chromatin organisation • DNA structure • nucleosomes • 30 nm fibre • nucleoid-associated proteins

11.1 Introduction

In all organisms DNA is maintained in a highly compacted state complexed with abundant basic proteins. The extent of this compaction can range from ~1,000-fold in the bacterial nucleoid to ~10,000-fold in eukaryotic metaphase chromosomes. In addition to the necessity for compaction the genetic specification function of DNA also requires that the appropriate encoded information be accessible for transcription.

A. Travers (✉)

Fondation Pierre-Gilles de Gennes pour la Recherche, c/o LBPA,
École Normale Supérieure de Cachan, 94235, Cachan, France;
MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 0QH, UK
e-mail: aat@mrc-lmb.cam.ac.uk

G. Muskhelishvili

Jacobs University, Campus Ring 1, D-28759, Bremen, Germany

The dual requirements of compaction and selective accessibility imply that the complex of DNA and abundant basic proteins, defined here generally as chromatin, must possess a high degree of structural organisation and that the regulation of transcription at the level of the gene may involve substantial structural transitions. In this article we summarise the current understanding of the organisation of eukaryotic chromatin and discuss the extent to which the organisational principles are also apparent in prokaryotic chromatin.

11.2 Chromatin-Associated Proteins

The most abundant basic proteins in the eukaryotic nucleus are the core histones. These are comprised predominantly of H2A, H2B, H3 and H4 which associate to form the histone octamer (Thomas and Kornberg 1975). In addition to these canonical core histones, variants of, especially, histones H2A and H3 occur at lower copy numbers. Amongst other abundant proteins are the linker histones, which aid the compaction of the chromatin fibre, and in some organisms may also occur in several forms (Thomas 1999; Izzo et al. 2008). Other proteins involved in determining chromatin structure and function include the heterogeneous category of HMG (High Mobility Group) proteins. These are themselves split into distinct functional classes – the HMGA, HMGB and HMGC proteins (Bustin 2001a). All these proteins interact directly with nucleosomes (Travers 2005). The HMGA and HMGN proteins antagonise linker histone binding, and the HMGB proteins increase the accessibility of wrapped nucleosomal DNA (Bustin 2001b; Ragab and Travers 2003). Importantly the HMGB proteins can also increase the torsional and axial flexibility of DNA by transiently introducing tight bends (Thomas and Travers 2001).

The DNA in the bacterial nucleoid is also organised by abundant basic proteins but with the exception of the SMC proteins there appear to be no direct paralogues of abundant eukaryotic nuclear proteins. Although some nucleoid-associated proteins (NAPs) such as HU are often described as ‘histone-like’ the only parallels are functional. For example HU and FIS can wrap DNA to a similar degree as the histone octamer (Maurer et al. 2009). Again HU can act as a functional counterpart of the eukaryotic HMGB proteins, although these proteins are structurally completely distinct. Thus HMGB proteins can in large part substitute for HU function in establishing *lac* repression loops and maintaining negative superhelicity and nucleoid condensation in vivo (Paull and Johnson 1995; Becker et al. 2005).

11.3 The Structural Organisation of Eukaryotic Chromatin

11.3.1 *The Organisation of Nucleosomal Arrays*

In the eukaryotic nucleus the DNA is compacted by 1,000–10,000 fold principally by interaction with the histone proteins. In vivo chromatin can assume different structural states with varying degrees of compaction but little is known about the

organisation of the most highly compacted forms. The fundamental structural unit of chromatin is the nucleosome core particle in which an octamer of histones wraps ~147 bp of DNA in ~1.65 superhelical turns (Luger et al. 1997). Both the accessibility of regulatory regions and the formation of higher-order structures depend on the positional placement of nucleosomes on the DNA sequences. While there is abundant evidence that nucleosomes can occupy preferred positions on nucleosomal DNA the nature of the factors that determine positioning is less apparent. Kornberg (1981) originally posed the question as to whether the location of nucleosomes in chromatin is specific or statistical. This proposition was subsequently refined to conclude that the observed distribution of nucleosomes in gene-specific arrays could be explained by statistical positioning associated with a boundary constraint for each array, such as a regulatory protein (Kornberg and Stryer 1988). Neglecting end effects statistical positioning implies that the potential number of histone octamer binding sites in a DNA sequence is essentially equivalent to the number of base-pairs in that sequence, whereas specific positioning implies that, at one extreme, the observed pattern of nucleosome positions is uniquely defined by the DNA sequence. This latter possibility is countered by the observation that the repeat length of nucleosomal DNA in higher eukaryotes is tissue dependent (Compton et al. 1976; Thomas and Thompson 1977). Recent data from genomic parallel sequencing strongly support a mode of positioning in which there are many potential histone octamer binding sites in any particular stretch of DNA sequence but that an array, which may be regarded as the unit of gene-specific chromatin structure, is specified by boundary constraints (Mavrich et al. 2008; Valouev et al. 2008). In particular Mavrich et al (2008) showed that in yeast the 5' nucleosome of an array was generally well positioned while the positions of subsequent nucleosomes were increasingly less well defined or 'fuzzy'. Additionally Fire and his collaborators (Valouev et al. 2008) showed that in the nematode *Caenorhaditis elegans* a diversity of nucleosome positions was observed at most loci although the degree of diversity varied between loci. A similar conclusion was drawn from analysis of nucleosome positions at the yeast CUP1-2 and HIS3 loci where multiple arrays with different occupancies were observed (Shen and Clark 2001; Kim et al. 2006).

It is also well established that binding of DNA to the histone octamer is strongly favoured both by DNA sequences with a preferred bending trajectory (Drew and Travers 1985; Satchwell et al. 1986; Travers and Klug 1987; Shrader and Crothers 1989; Piña et al. 1990b; Olson et al. 1998; Widom 2001) and by DNA flexibility (Virstedt et al. 2004).

This is supported by two lines of evidence. In vitro intrinsically bent DNA preferentially binds octamers (Hsieh and Griffith 1988; Wolffe and Drew 1989). Additionally the sequence organisation of DNA isolated from chicken erythrocyte and yeast core nucleosome particles demonstrates a periodic distribution of AA/TT dinucleotides (reflecting the pattern of the trinucleotide AAA/TTT and not that of the isolated dinucleotide) and a similar distribution of GC dinucleotides in the opposite phase (Satchwell et al. 1986; Widom 2001). These patterns confer anisotropic bendability on the DNA in general (Travers and Klug 1987) and thus define a preferred direction of bending. They do not however by themselves specify translational positioning, that is, where an octamer is placed relative to the DNA sequence. A distinct

feature of the periodic AA/TT pattern is a change to the opposite phase in the vicinity of the midpoint or presumed structural dyad (Satchwell et al. 1986), leading to the proposal that this discontinuity could act as a translational positioning signal (Travers 1987). This proposition is supported by the finding that ~72% of yeast in vivo nucleosome positions mapped by partial MNase digestion have a periodicity pattern consistent with this feature (Caserta et al. 2009). These same nucleosomes contain a region of coherently bendable DNA on average 10–40 bp from the midpoint on one

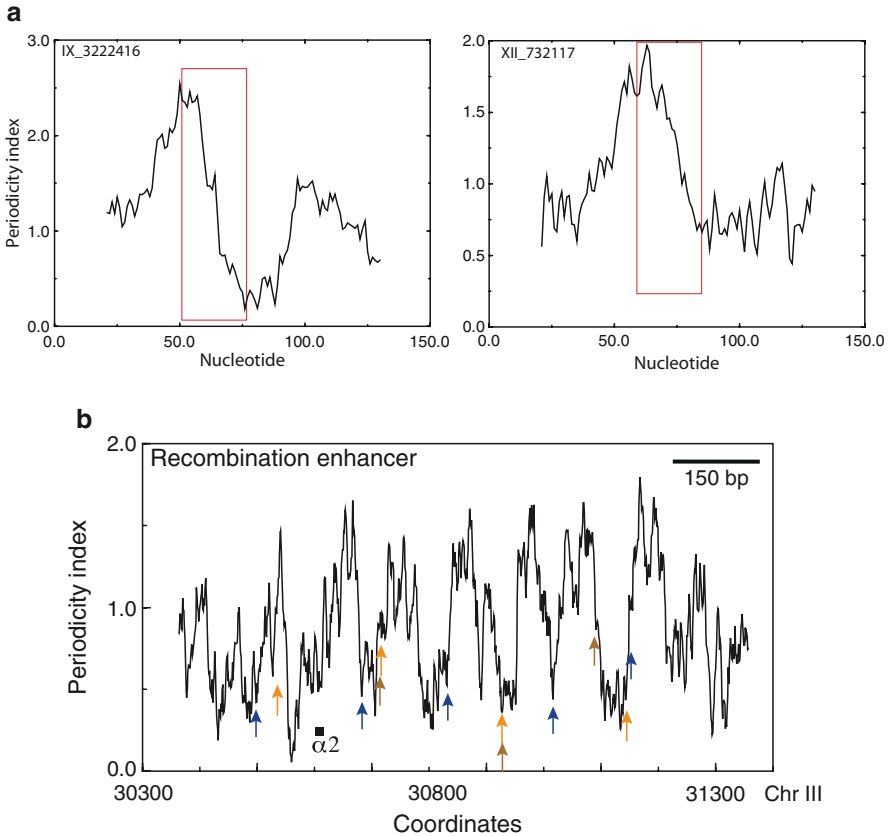


Fig. 11.1 (a) Examples of a periodicity signature for nucleosome positioning from yeast. The periodicity index is a measure of the periodic occurrence of (AA,AT,TA,TT) and (CC,CG,GC,GG) in opposite phases, corrected for amplitude as determined in the Satchwell et al. (1986) sequence compilation. For comparison a designed rotational positioning sequence has an approximately uniform periodicity index of ~2.7 (Caserta et al. 2009). The boxed region covers the change from a low to a high region of periodicity in the approximate centre of the mapped positions. (b) Comparison of partial MNase mapping data (Weiss and Simpson 1997) with that of parallel sequencing (Mavrich et al. 2008) and tiling array (Lee et al. 2007) analysis at the recombination enhancer locus. The resolution of the techniques is variable but note that the partial MNase mapping identifies two apparent positioned nucleosomes which are not identified by the other techniques. Blue arrows, positions identified by partial MNase mapping; orange arrow, positions identified by parallel sequencing, brown arrows, positioned identified by tiled microarray (Reproduced with permission from Travers et al. 2009, submitted)

or both sides (Fig. 11.1a). In contrast, an artificial positioning sequencing consisting only of tandem 20 bp repeats (Shrader and Crothers 1989) that fails to position nucleosomes translationally in vivo (Tanaka et al. 1992; Wallrath et al. 1994) lacks a significant change of AA/TT phase at the midpoint. Furthermore telomeric DNA sequences which, in vitro, have a low affinity for the histone octamer and poorly position octamers (Pisano et al. 2006) possess a uniformly low bendability. These data imply that many nucleosome positions correlate with a specific sequence pattern that determines the structural properties of DNA. This minimum length of this pattern is about only 30–50 bp suggesting that it constitutes a half-site for the binding of the H3/H4 tetramer during nucleosome formation.

Two other sets of observations are relevant to this scenario. First, the number of potential positions available on genomic DNA, as defined by the bendability pattern of nucleosomal DNA is $\sim 3\text{--}4\times$ greater than the actual number of nucleosomes (Caserta et al. 2009). However in one case, that of the MOX promoter (Costanzo et al. 1995), the positions of two mutually exclusive arrays mapped in vivo under different environmental conditions, both correspond to positional signals in the DNA (Caserta et al. 2009). In addition the amplitude of the positioning signals varies between genes. For example the average bendability of the DNA of the CUP1-2 and HIS3 genes is low, consistent with the multiplicity of mutually exclusive arrays observed on these genes (Shen and Clark 2001; Kim et al. 2006). In contrast, DNA which directs well-defined nucleosome arrays, for example the mating type recombination enhancer and the SAC7 gene (Weiss and Simpson 1997; Caserta et al. 2009), is characterised by a succession of strong putative positioning signals (Fig. 11.1b). Most pertinent is the observation that the DNA signature for translational positioning is a characteristic of nucleosomes in the vicinity of a promoter and not of non-promoter nucleosomes (Caserta et al. 2009). These observations argue that observed preferred nucleosome arrays may indeed require a boundary constraint but that the associated positioning is neither strictly statistical nor necessarily uniquely defined. It follows that the documented strongly positioned nucleosome arrays are likely aligned by an ‘organiser’ acting in concert with nucleosome remodelling complexes (Caserta et al. 2009). Such an organiser could be a strong intrinsic positioning signal specified by the DNA sequence or alternatively could be a sequence-specific DNA-binding protein, or indeed both (Fig. 11.2). Potential examples of both mechanisms have previously been described. A strong intrinsically positioned nucleosome is found in the MMTV LTR promoter (Richard-Foy and Hager 1987; Piña et al. 1990a) while the binding of the yeast $\alpha 2$ repressor correlates with aligned nucleosome arrays occluding the TATA boxes of repressed genes (Shimizu et al. 1991).

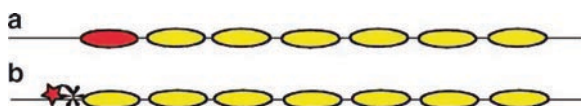


Fig. 11.2 Possible mechanisms for organising a nucleosome array. (a) A strong positioning sequence accurately specifies the position of the 5' nucleosome in the array. (b) A sequence-specific DNA-binding protein contacts and stabilises the 5' nucleosome. In practice both mechanisms probably operate

Second, the sequence pattern of the DNA from positioned nucleosomes identified by partial MNase digestion differs significantly from that identified by methods which involve digestion to mononucleosomes. In particular the average A/T content at the midpoint of the partial MNase set is 72% while the total average A/T composition (62%) is essentially identical to that of yeast genomic DNA (61.7%) (Caserta et al. 2009). By comparison, the midpoint A/T content of the set of core nucleosome DNA sequences isolated from yeast by H2A.Z tagging and subsequent pyrosequencing is 61% (Albert et al. 2007) and that of a yeast set determined by sequencing the DNA of isolated core nucleosomes is ~59% (Segal et al. 2006). These figures imply a selective loss of nucleosomal DNA with TA and/or AA/TT at its midpoint during limit digestion by micrococcal nuclease to mononucleosomes. Indeed, overall depletion of the TA dinucleotide was noted for the set of chicken core particle sequences (Satchwell et al. 1986). Such an effect could bias any recovered set of nucleosomal DNA sequences and result in an artificially low A/T content within the set. This could explain why nucleosome positions mapped by partial MNase digestion do not always correspond to those mapped by other methods requiring limit digestion. Nevertheless the average sequence organisation of nucleosomal DNA sequences recovered from isolated nucleosomes is entirely consistent with the structure of the core nucleosome (Luger et al. 1997) and therefore these sequences likely represent *bona fide* octamer binding sites. Given that in vitro the octamer can bind to a large number of overlapping sites with varying affinities (Gencheva et al. 2006), the modest correspondence between sequences identified by partial and limit MNase digestion may reflect overlapping sets of positions rather than erroneous mapping, again supporting the notion of a fluid chromatin structure that can accommodate multiple array settings. However, importantly any selection of nucleosomes by limit MNase digestion would compromise estimates of nucleosome occupancy using that method.

11.3.2 The 30 nm Fibre

The 30 nm fibre is the paradigm for the first order of folding of a nucleosome array, yet its relevance to functional chromatin in vivo is, even now, not firmly established (Eltsov et al. 2008). The broad outlines of the structure of this fibre, apart from the trajectory of the linker DNA (Robinson and Rhodes 2006; Wu et al. 2007), are well established. Typically its basic architecture is a helical array of nucleosomes with a diameter of 30–32 nm that is essentially independent of linker lengths in the range of 20–60 bp, where linker length is defined as the number of base-pairs between adjacent nucleosome core particles (Butler and Thomas 1980; Thomas and Butler 1980; Worcel et al. 1981; Bates et al. 1981; Pearson et al. 1983; Widom and Klug 1985; Widom 1986). Most experimental determinations of the mass/unit length of both the native and reconstituted 30 nm fibre is approximately 6–7 nucleosomes/11 nm in solution (Thoma et al. 1979; Widom and Klug 1985; Felsenfeld and McGhee 1986; Gerchman and Ramakrishnan 1987; Ghirlando et al. 2004; Hizume et al. 2005; Ghirlando and Felsenfeld 2008; Kruithof et al. 2009). Although a mass/unit length

of ~ 6 nucleosomes/11 nm is a robust experimental value for most 30 nm fibres. Higher packing densities of 13 ± 2 nucleosomes/11 nm have been reported for chromatin from sea cucumber sperm (Athey et al. 1990). In addition, electron microscopy revealed that artificial nucleosome arrays of a strong positioning sequence with constant and optimal linker lengths ranging from 10–40 bp condense to fibres with a mass/unit length of $\sim 11 \pm 1$ nucleosomes/11 nm while fibres with longer linkers of 50–70 bp can assume an even higher packing density of 14–15 nucleosomes/11 nm (Robinson et al. 2006).

A major contentious question concerning the structure of the 30 nm fibre is whether it is organised as a 1-start stack of nucleosomes or solenoid, essentially a single continuous helix or, as a 2-start helix, or zigzag, with two continuous stacks, analogous to the structure of DNA (Fig. 11.3). All current models of the structure of the 30 nm fibre assume, either implicitly or explicitly, that all experimental

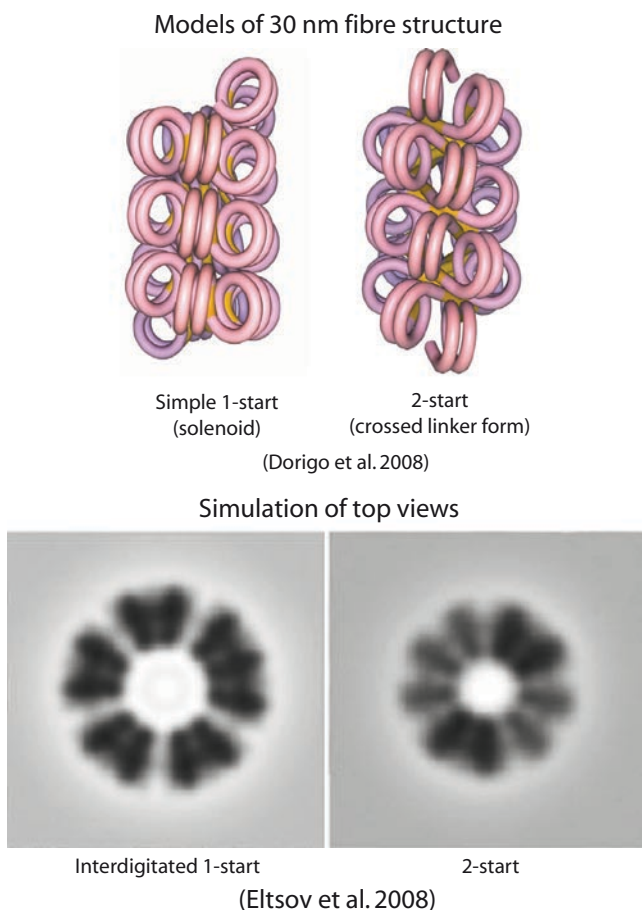


Fig. 11.3 Models for 1-start and 2-start topology of the 30 nm fibre (Reproduced with permission from Eltsov et al. 2008 and Dorigo et al. 2004)

observations pertain to essentially the same structure. The models divide easily into two classes – the solenoid class in which a 1-start helical stack of nucleosomes is linked together by bent DNA between adjacent octamers (Finch and Klug 1976; Thoma et al. 1979; Robinson et al. 2006) and the 2-start zigzag class in which, in the simplest examples, two helical stacks of nucleosomes are connected by a relatively straight DNA linker. The latter class further divides into the helical/twisted ribbon model in which the linker DNA is oriented at angles varying from 0–50° to the fibre axis i.e. along the length of the fibre (Worcel et al. 1981; Woodcock et al. 1984) and the crossed-linker model in which the linker DNA is oriented approximately perpendicular to the fibre axis, i.e. across the fibre (Williams et al. 1986; Schalch et al. 2005).

There is now compelling experimental evidence that the ‘30 nm’ fibre can adopt a 2-start organisation. The crystal structure of a tetranucleosome with a 20 bp linker but without linker histone, revealed two stacks of nucleosomes connected by ‘straight’ linkers (Schalch et al. 2005). Chromatin fibres formed from a longer array of the same positioning sequence also adopt the same geometry (Routh et al. 2008). Additionally cross-linking studies on longer reconstituted arrays with uniform optimal linkers indicated that nucleosome i was immediately adjacent to nucleosomes $i \pm 2$, as predicted by a 2-start helix (Dorigo et al. 2004). Crucially the same geometry was observed for regular arrays with linker lengths 0, 20 and 40 bp when compacted either with divalent cations or linker histone. These direct demonstrations of 2-start connectivity are consistent with other observations – the DNA cleavage patterns induced by ionising radiation in intact cells (Rydberg et al. 1998), the ‘straight’ trajectory of linkers in a slightly extended fibre and their relative resistance to UV-induced pyrimidine dimer formation (Pehrson 1995). Nevertheless, the dependence of the diameter of compact fibres cannot be explained by either of the crossed-linker or helical ribbon models alone (Robinson et al. 2006; Wu et al. 2007). Several possible models have been proposed to account for this apparent anomaly. These include a 4–5 start helical structure (equivalent to an interdigitated 1-start helix) (Robinson et al. 2006; Kepper et al. 2008; Routh et al. 2008), a 3-start helix (Staynov and Proykova 2008) and a model in which the connectivity, and hence the number of coiled nucleosome stacks, depends on linker length (Wong et al. 2007). However if it is assumed that the helical ribbon and crossed-linker structures are topologically equivalent the experimental observations can be simply explained by assuming a smooth linker-dependent transition from one form to the other (Wu et al. 2007). More unequivocally, the model proposed by Robinson et al. (2006) for a compact fibre based on a 177 bp nucleosome repeat length predicts that this fibre adopts a 5-start helical configuration (Eltsov et al. 2008) whereas class-average cryo-em images of a view down the helical axis of a compact 34 nm diameter fibre are only consistent with a 2-start configuration (Robinson 2005). Recently single molecule experiments showing that a folded array of nucleosomes lacking linker histone stretched from 50 to ~160 nm under 4 pN tension was interpreted as favouring a solenoid model (Kruithof et al. 2009). However this interpretation assumes that a single nucleosome stack of ~300 nm compacted to 160 nm under tension. An alternative interpretation would be that two intertwined nucleosome stacks of ~150 nm

length, would unwind and both slightly extend to ~160 nm each under tension, consistent with a 2-start helix.

The topological connection between the compaction of the fibre and changes in DNA wrapping is described by the ‘linking number paradox’, first enunciated by Finch et al. (1976) in the context of a mismatch between the number of negative superhelical turns of DNA wrapped in a nucleosome core particle (1.5–2) and the experimentally observed value of $\Delta L = -1$ on nucleosome formation (Germond et al. 1976; Keller et al. 1978). Finch et al. (1976) suggested that this mismatch could be resolved if the DNA wrapped on the surface of the histone octamer adopted a helical repeat of 10 bp/turn instead of the 10.4 bp/turn characteristic of DNA in solution (Peck and Wang 1981). However, subsequent analysis of nucleosomal core DNA sequences showed that the actual helical repeat of octamer-bound DNA is ~10.2–10.25 bp/turn (Drew and Travers 1985; Satchwell et al. 1986), a value that cannot account fully for the paradox (Travers and Klug 1987). The discrepancy between observed and calculated values arises because the topology of a nucleosome array – whether folded or unfolded – is not fully described by that of its component core particles, as visualised in the crystal structure, since both the trajectories of the exiting and entering DNA and the intrinsic helicity of the fibre also contribute. For a tetranucleosomal array lacking linker histone the magnitude of the intrinsic helicity is given by the relative coiling of -71.5° of the two nucleosome stacks (Schalch et al. 2005). This negative coiling in turn compensates for a positively-coiled trajectory of the exiting and entering DNA, i.e. in the opposite sense to the wrapped DNA (Zivanovic et al. 1988; Travers 2009). The magnitude of this positive component is respectively increased or decreased, relative to the nucleosome core particle, by linker histones or histone acetylation respectively (for a formal proof see Travers (2009)). Since the positive component is directly coupled to the extent of wrapping so also is the compensating coiling and resultant compaction. This solution to the linking number paradox has two important implications. First, each nucleosome in an array can be considered to contribute directly to the compaction of the fibre and second, the processes of coiling and uncoiling, to a first approximation, approach topological neutrality. In other words during condensation and decondensation the generation of torque is minimised, but not necessarily absent. In addition this solution constrains the chromatin fibre to 2-start connectivity.

These topological considerations allow a simple description of the folding and unfolding of the chromatin fibre (Fig. 11.4). In essence the fibre acts as a tunable coil. In agreement with the scattering measurements of Bordas et al. (1986) at low ionic strength the fibre exists as a loose, high-pitch helical structure. With increasing ionic strength or addition of linker histone the internucleosomal distance decreases with a concomitant decrease in the pitch of the 2-start helix. This process would likely be highly cooperative and rapid. Decondensation is simply the reverse of condensation. The degree of condensation of the fibre is tunable and dependent on such factors as the regularity of nucleosome spacing, histone modifications and the nature of the linker histone. Since the number of nucleosomes/helical turn of a fibre can vary (Robinson et al. 2006; Athey et al. 1990), the degree of compaction and the form of the fibre may depend on this parameter. For example, whereas the

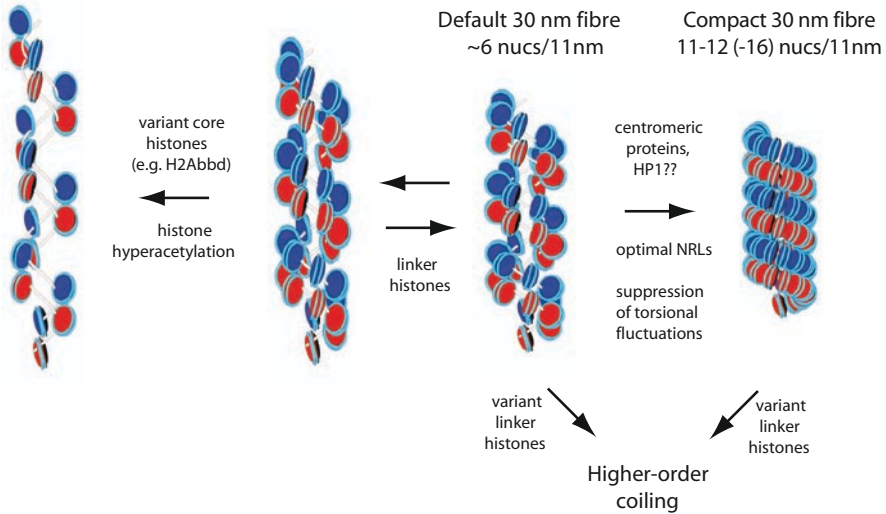


Fig. 11.4 The folding and unfolding of the chromatin fibre. The fibre extends or compacts by a concertina mechanism. Different states of the fibre are stabilised by different proteins (Adapted from Bassett et al. 2009)

canonical 30 nm fibre contains ~6 nucleosomes/11 nm (Gerchman and Ramakrishnan 1987; Ghirlando et al. 2004; Ghirlando and Felsenfeld 2008) those reconstituted on arrays of optimally spaced nucleosome positioning signals *in vitro* may contain between 11 and 16 nucleosomes/11 nm (Robinson et al. 2006). Thus per turn there will be a greater contribution to coiling from the nucleosomes for the more compact relative to the canonical fibres. This is consistent with the interpretation that the more compact fibres can, for short linker lengths, assume a more coiled form of the fibre (see above). In addition, the individual contributions of nucleosomes to coiling could also, in principle, facilitate the formation of higher-order structures beyond the 30 nm fibre. If coiling transitions were not completely topologically neutral such that, for example, the coiling in the 30 nm fibre structure is less than the total topological nucleosomal contributions, then any imbalance could be absorbed by a further coiling of the 30 nm fibre on itself, either as a plectoneme or as a toroid. Importantly since these structures would contain many nucleosomes the actual contribution per nucleosome to the higher-order coiling would be small relative to that driving the coiling of the 30 nm fibre. Nevertheless, although higher-order coiling is an attractive solution to the structural organization of highly condensed chromatin and is supported by electron microscopy (Urata et al. 1995; Kireeva et al. 2004) the actual structures involved remain to be established experimentally. Indeed recent visualization of metaphase chromosomes by cryo-electron microscopy failed to identify any structures resembling 30-nm fibres but instead the chromatin fibre appeared in a highly disordered and interdigitated state, which on the local scale may be compared to a polymer melt (Eltsov et al. 2008).

11.3.3 Organization and Topology of Bacterial Chromatin

The organisation of eukaryotic chromatin, at least for nucleosome arrays and the 30 nm fibre, may be considered as successive degrees of coiling. Thus the DNA ‘rope’ is wrapped as a toroidal coil on the histone octamer and the preferred position of the octamer on DNA is, at least in part, determined by its flexibility and bendability. Because there are more potential octamer binding positions than nucleosomes the setting of a nucleosome array must be ‘nucleated’. The most prominent positioning signals generally occur at the 5′ end of an array (Caserta et al. 2009) suggesting that this nucleation occurs usually occurs at one end. The array itself can then coil to become the 30 nm fibre and possibly then the 30 nm fibre acts as a ‘rope’ which can be further coiled.

To what extent are these principles also apparent in prokaryotic chromatin? Many of the NAPs, including HU, H-NS, FIS and LRP, but possibly not IHF and Dps constrain negative supercoils (Schneider et al. 2001; Travers and Muskhelishvili 2005a,b; Maurer et al. 2006). HU (Rouvière-Yaniv et al. 1979; Kobryn et al. 1999; Swinger et al. 2003; Guo and Adhya 2007; Maurer et al. 2009), FIS (Maurer et al. 2006, 2009) and possibly LRP (Pul et al. 2007) can wrap DNA as a toroid while H-NS can stabilise the plectonemic form of supercoiled DNA (Schneider et al. 2001; Lang et al. 2007; Maurer et al. 2009) and thereby establish long-range structures repressing transcription at specific loci. But how are such structures established? In exponentially growing *Escherichia coli* the total concentration of NAPs is ~200 μM (Muskhelishvili and Travers 2009) while that of the most abundant NAPs is normally in the range 10–80 μM , implying a concentration of the unbound protein at least in the range of 0.1–1 μM . Like the histone octamer HU, H-NS (Tupper et al. 1994) and FIS (Lazarus and Travers 1993; Bétermier et al. 1994) can bind to a wide range of DNA sequences with varying affinities and thus in vivo are likely to compete for binding to many sites on the DNA.

The establishment of structures at specific loci appears to require two components, a high affinity binding site or sites which nucleate subsequent cooperative binding. This mechanism provides selectivity. One such example is the binding of FIS to the UAS of the *tyrT* promoter. The UAS contains three high affinity FIS binding sites, two of which have K_d values of 7.5–30 nm in vitro (Lazarus and Travers 1993). Binding to the three sites by FIS alone is highly cooperative (Pemberton et al. 2002) and is also cooperative with RNA polymerase binding at the core promoter (Muskhelishvili et al. 1995). The K_d values for the tightest sites are in this case >1,000-fold lower than the total concentration of FIS in the cell. Another example of nucleation on a more extensive scale is the repression of the *E. coli proU* locus by H-NS. The NRE (Negative Repression Element) contains two identical sites that bind H-NS with high affinity in vitro (K_d ~15 nM), again much lower than the total concentration of this NAP in the cell (Muskhelishvili and Travers 2009). A single such site can nucleate cooperative binding (Bouffartigues et al. 2007). This cooperative binding might be to immediately adjacent DNA sequences or to a more distant DNA duplex by bridging, or to both

generating a plectonemic structure (Dame et al. 2000; Schneider et al. 2001; Lang et al. 2007; Maurer et al. 2009). Nevertheless, the spreading of H-NS is also probably dependent on DNA sequence. High affinity binding sites for H-NS are found in many transcription units repressed by H-NS (Lang et al. 2007). These sites, as in *proU*, are flanked by A/T rich regions which are preferred binding sites for higher H-NS concentrations in vitro and in vivo (Navarre et al. 2006). These sites would thus facilitate spreading and would be analogous to the octamer positioning sites occurring within arrays at yeast loci. Again as in the case of the histone octamer the precision of H-NS binding at these sites may be less than that at the primary nucleating site. In the case of long transcription units, for example that encoding the genes for flagella biogenesis, sequences with high homology to high-affinity binding sites occur at irregular intervals throughout the transcription unit (Lang et al. 2007) suggesting that maintenance of the repressive structure may require reinforcement. Nevertheless In essence the principle of producing a relatively stable three-dimensional chromatin structure is the same although the nature of the structure may be different in prokaryotic and eukaryotic chromatin.

Although the coiled structures formed by NAPs individually are beginning to be become apparent (Maurer et al. 2009), as in eukaryotes little is known about the structures involved in higher orders of condensation that would be necessary to form the compact nucleoid. This is especially true in stationary phase cells where the NAP Dps condenses the DNA into a structure that has the properties of a liquid polymer (Wolf et al. 1999).

11.4 The Regulation of Chromatin Structure

11.4.1 *Structural Transitions in Eukaryotic Chromatin*

An important consequence of the structural organisation of chromatin in the nucleus and nucleoid is that the regulation of chromatin accessibility and hence of gene expression also involves substantial changes in chromatin structure. In both bacteria and eukaryotes the DNA is negatively supercoiled but whereas in eukaryotes the total overall superhelicity remains fairly constant, in prokaryotes it can vary substantially (reviewed in Dorman 1996; Travers and Muskhelishvili 2005b, 2007). One consequence of this difference is that whereas in prokaryotes different basic proteins constrain different superhelical densities of DNA and local DNA trajectories, in eukaryotes changes in chromatin accessibility depend on the regulation of the folding and unfolding of the chromatin fibre regulating the transitions between the most open and the most highly condensed fibres.

Since a very high proportion of eukaryotic DNA is packaged in nucleosomes, which by themselves possess very limited intrinsic mobility, structural transitions in chromatin are required for changes in gene regulation and accompany the processes of transcription, DNA replication, DNA repair and recombination.

Structural transitions which affect the state of chromatin, for example, euchromatic or heterochromatic, generally alter the equilibrium between folded and unfolded states and may be effected by histone modifications, variant core histones and variant linker histones cooperating with ancilliary proteins such as heterochromatin protein 1 (HP1).

The facile folding and unfolding of the 30 nm fibre in which the degree of coiling of the fibre and hence of compaction is directly related to the extent of DNA wrapping around the histone octamer (Travers 2009) can, in principle, be modulated by varying this parameter. For example, the histone variant H2Abbd, which is normally associated with transcriptionally active chromatin (Gautier et al. 2004) supports less DNA wrapping than the canonical H2A and is associated with an unfolding of the fibre (Bao et al. 2004; Doyen et al. 2006; Zhou et al. 2007). It is also possible, but not demonstrated, that different linker histones could induce different degrees of coiling and hence of compaction. In principle, as described above, a linker histone which induced more coiling than could normally be accommodated by a 30 nm fibre would promote the coiling of the 30 nm fibre upon itself to produce a more compacted higher-order structure.

The existence of different modes of nucleosome packing within fibres implies that the nucleosome–nucleosome contacts in the different structures may vary. In particular the contacts between adjacent nucleosomes, which are dependent on histone tails (Dorigo et al. 2003; Bertin et al. 2004; Hizume et al. 2009), may be mediated by different tail interactions depending on the degree of compaction. This implies that the degree of compaction may itself depend structurally on the nature of histone modifications.

Histone acetylation is known to be a potent regulator of overall chromatin compaction. When cells are treated with trichostatin A, an inhibitor of histone deacetylases, all the previously heterochromatic regions decondense to form structures that have the appearance of euchromatin (Tóth et al. 2004). While such a decondensation could be an indirect consequence of effector proteins such as chromatin remodelling factors, it may also be the result of the abrogation of positive charge on lysine acetylation, which would likely affect nucleosome–nucleosome interactions (Durrin et al. 1992; Dorigo et al. 2003; Shogren-Knaak et al. 2006; Robinson et al. 2008) and also reduce DNA wrapping (Norton et al. 1989, 1990; Bauer et al. 1994). The observed effects of histone hyperacetylation on the folding of the chromatin fibre *in vitro* are variable (McGhee et al. 1983; Tse et al. 1998). Notably the largest effects of acetylation on chromatin folding have been observed with reconstituted arrays with a constant and nearly optimal linker length (Tse et al. 1998; Shogren-Knaak et al. 2006).

The histone H4 tail is critical for the formation of densely packed chromatin states (Shogren-Knaak et al. 2006; Robinson et al. 2008), and appears to interact with an acidic patch present on the H2A/B dimer of a neighbouring nucleosome in the fibre (Luger et al. 1997). Acetylation of H4 K16 antagonises such an interaction in both long and short fibres (Robinson et al. 2007), and methylation of the nearby K20 has been shown to encourage folding of the fibre by shifting the equilibrium from a less dense (40S) to highly condensed (55S) array (Nishioka et al. 2002). This may be mediated structurally by allowing interaction of the neighbouring R19

residue with a neighbouring nucleosome (Lu et al. 2008), and in vivo by antagonising acetylation of K16. Mutagenesis of the acidic patch on the H2A molecule or replacement with the histone variant H2Abbd, which lacks the acidic region as well as decreasing DNA wrapping, also shifts the equilibrium towards a less dense (40S) array, that is more conducive to transcription, further supporting this hypothesis (Zhou et al. 2007).

Histone tails may also play a role in “higher order” folding beyond the 30 nm fibre by mediating internucleosomal contacts between arrays (Kan et al. 2009). Lysine to glutamate substitutions in histone tails, aiming to mimic the charge neutralisation upon acetylation show that the effects are site-specific. H2B and H4 mutants inhibit array self-association, as measured by aggregation of chromatin fibres at increasing magnesium concentrations, such that if both are mutated, aggregation is completely prevented (Kan et al. 2009). Not all acetylation is equivalent, since mimics of H3 acetylation appear to have a stronger effect on individual nucleosome structure than on higher order folding (Wang and Hayes 2008).

11.4.2 *Structural Transitions in Bacterial Chromatin*

A major difference between bacteria and eukaryotes is that in the former, gene expression is more immediately regulated by changes in external environmental conditions, such as osmolarity and nutrient availability. (One exception to this would be the stress response induced by heat shock.) Thus, in Proteobacteria both the composition of the nucleoid and the overall negative superhelical density vary substantially with the growth cycle (Dorman 1996; Travers and Muskhelishvili 2005a), such that in early exponential phase FIS and HU α , whose bound DNA in both cases has a high local superhelical density, predominate while in late stationary phase Dps, which constrains negligible superhelicity, is the major protein. These changes correlate with a transition from high to low negative supercoiling levels. In contrast the total levels of H-NS and its paralogues appear to remain relatively unperturbed.

During the exponential phase of growth several inducible systems – including the operons including the genes specifying flagella and pili as well as the osmotically regulated *proU* operon are repressed by H-NS but when activated require HU for optimal activity (Oberto et al. 2009; Berger M, Geertz M, Faracs A, Zhelyazkova P, Brix K, Travers A, Muskhelishvili G., unpublished observations). A similar pattern is observed in a *hupA* mutant strain which constitutively expresses genes normally repressed by H-NS (Kar et al. 2005). Since H-NS and HU constrain plectonemic and toroidal structures respectively this implies a change in the form of superhelical constraint. Expression of the *proU* operon is induced by osmotic shock (Cairney et al. 1985) which increases the negative superhelical density in vivo (Higgins et al. 1988; Cheung et al. 2003; Gralla and Vargas 2006). This correlates with the relief of H-NS repression, an effect which is consistent with an observed decrease in affinity for the *proU* nucleating sites on highly supercoiled

DNA in vitro (Bouffartigues et al. 2007). However high levels of *proU* transcription are promoted by HU (Oberto et al. 2009; Berger M, Geertz M, Faracs A, Zhelyazkova P, Brix K, Travers A, Muskhelishvili G., unpublished observations). Such a decrease in H-NS affinity could also change the relative competitive advantage of HU and H-NS binding to the same stretch of DNA implying that the repressed and active forms of *proU* chromatin are plectonemic and toroidal respectively (Fig. 11.5a). In both structures the DNA loop that contains the promoter regions would be conserved although the exit and entry trajectories of the DNA could differ.

Another example of a structural transition is the induction of operons such as *lac*, *araBAD* and *gal* where the inactive state is maintained by a repression loop stabilised by HU (Krämer et al. 1987; Bellomy et al. 1988; Lee and Schleif 1989; Aki et al. 1996; Becker et al. 2005). The respective repressors, LacR, AraC and GalR are all bivalent DNA binding proteins which bind to bent DNA and contain a flexible hinge region between the DNA-binding domains. However while the LacR DNA binding domains bind to the outside of the bend (Lewis et al. 1996), that of AraC, which has sequence similarities to GalR, binds to the inside of the bend (Saviola et al. 1998). Although the lengths of repression loops are variable, for *lac* and *araBAD* the optimal size is 70–90 bp, a length which corresponds well to the dimensions of the toroidal fibre stabilised by HU in vitro. Such a loop could thus be regarded as a single turn of a toroidal fibre stabilised by a bivalent repressor trapping the DNA of successive turns induced by HU binding (Fig. 11.5b). In principle if HU acted as a DNA chaperone, its interaction with HU need only be transient (Travers et al. 1994). The three-dimensional organisation of the repressed state is thus different from that of H-NS-mediated silencing. Release of repression and the subsequent binding of polymerase could result in the formation of a longer loop with a concomitant repartitioning of writhe and twist.

11.5 Conclusions

We have argued that in both the prokaryotes and eukaryotes the packaging and utilisation of genetic material depends on the topological properties of DNA. One clue to the observed differences is provided by consideration of the spatiotemporal context in which these topological transitions take place. To a first approximation the human genome exceeds that of *E. coli* by 2,000-fold and has to be 10 times more compacted in size. Indeed, whereas 300 kb would be sufficient to encode 300 genes in *E. coli* in the human genome 300 genes can comprise 30 Mb. Furthermore, at least in some terminally differentiated eukaryotic cells particular genomic regions need to be silenced for years, while in eubacteria with a life span of less than an hour there is no such requirement. This implies that in general the transition from silenced to active configuration in prokaryotes needs be more facile.

The toroidal coiling of DNA in nucleosomes leads to a higher level of compaction compared to prokaryotic “half-nucleosomes” or the plectonemic coils and so

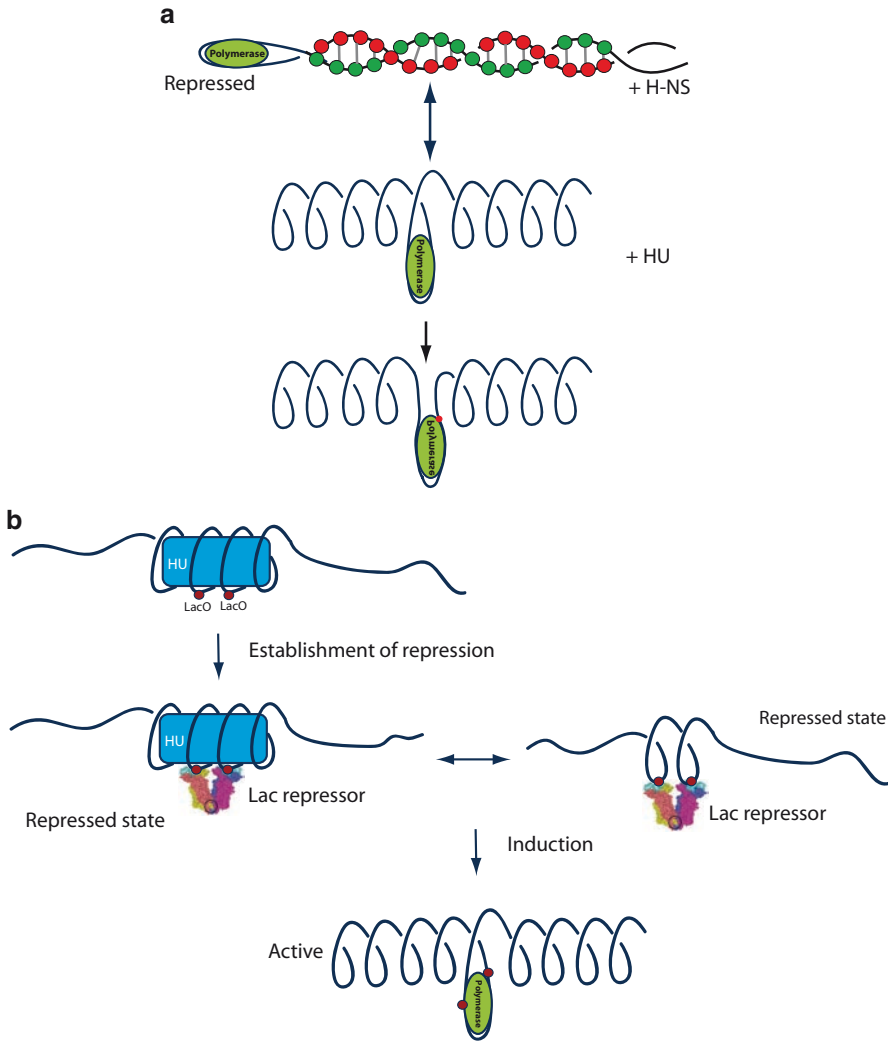


Fig. 11.5 Models for the repression and activation of transcription in the bacterial nucleoid. **(a)** Repression of *E. coli proU* operon by H-NS and its activation by HU resulting in a transition from plectonemic to toroidal coiling. This transition allows binding of RNA polymerase which can then utilise the energy of supercoiling to facilitate the opening of the transcription bubble. **(b)** Repression of the *lac* promoter by LacR. The repression loop is part of a toroidal coil transiently stabilised by HU. Induction results in the release of repressor and binding of RNA polymerase. *Lac* operator sites are indicated by red circles

is a preferential mode of packing. Indeed, as we have argued above, in eukaryotes changes in chromatin accessibility depend on the regulation of the folding and unfolding of the chromatin fibre by smooth transitions between the most open and the most highly condensed fibres. By contrast, in prokaryotes the same change

apparently involves a non-smooth transition between the plectonemic and toroidal shapes of DNA. Thus, stabilisation of distinct activity states of chromatin is achieved here by a higher energetic penalty associated with non-smooth topological transitions (Boles et al. 1990; Bliska et al. 1991). The flat and rigid plectonemic filaments would provide for reasonable silencing on the scale of a bacterial genome as opposed to the requirement of highly condensed fibres serving the same purpose in large eukaryotic genomes. Yet, in bacteria transitions between plectonemic and toroidal forms can be greatly facilitated by compositional changes of nucleoid-associated proteins.

In other words, the plectonemic fibres would be the favoured form of “silenced” chromatin in organisms with small genomes and short life spans, whereas in organisms with large genomes imposing high compaction requirements and having extended life spans highly condensed toroidal fibres would be favoured. We note that these latter, by virtue of the convertibility of plectonemic and toroidal coils in an elastic fibre, can adopt plectonemic coiling at a higher level of organizational complexity.

Acknowledgements We thank Dr. Malcolm Buckle for incisive comments.

References

- Aki T, Choy HE, Adhya S (1996) Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of *gal* transcription by GAL repressor. *Genes Cells* 1:179–188
- Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF (2007) Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* 446:572–576
- Athey BD, Smith MF, Rankert DA, Williams SP, Langmore JP (1990) The diameters of frozen-hydrated chromatin fibers increase with DNA linker length: evidence in support of variable diameter models for chromatin. *J Cell Biol* 111:795–806
- Bao Y, Konesky K, Park YJ, Rosu S, Dyer PN, Rangasamy D, Tremethick DJ, Laybourn PJ, Luger K (2004) Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *EMBO J* 23:3314–3324
- Bassett A, Cooper S, Wu C, Travers A (2009) The folding and unfolding of eukaryotic chromatin. *Curr Opin Genet Dev* 19:159–165
- Bates DL, Butler PJG, Pearson EC, Thomas JO (1981) Stability of the higher-order structure of chicken-erythrocyte chromatin in solution. *Eur J Biochem* 119:469–476
- Bauer WR, Hayes JJ, White JH, Wolffe AP (1994) Nucleosome structural changes due to acetylation. *J Mol Biol* 236:685–690
- Becker NA, Kahn JD, Maher LJ 3rd (2005) Effects of nucleoid proteins on DNA repression loop formation in *Escherichia coli*. *J Mol Biol* 349:716–730
- Bellomy GR, Mossing MC, Record MT Jr (1988) Physical properties of DNA in vivo as probed by the length dependence of the *lac* operator looping process. *Biochemistry* 27:3900–3906
- Bertin A, Leforestier A, Durand D, Livolant F (2004) Role of histone tails in the conformation and interactions of nucleosome core particles. *Biochemistry* 43:4773–4780
- Bétermier M, Galas DJ, Chandler M (1994) Interaction of Fis protein with DNA: bending and specificity of binding. *Biochimie* 76:958–967

- Bliska JB, Benjamin HW, Cozzarelli NR (1991) Mechanism of Tn3 resolvase recombination in vivo. *J Biol Chem* 266:2041–2047
- Boles TC, White JH, Cozzarelli NR (1990) Structure of plectonemically supercoiled DNA. *J Mol Biol* 213:931–951
- Bordas J, Perez-Grau L, Koch MH, Vega MC, Nave C (1986) The superstructure of chromatin and its condensation mechanism I. Synchrotron radiation X-ray scattering results. *Eur Biophys J* 13:157–173
- Bouffartigues E, Buckle M, Badaut C, Travers A, Rimsky S (2007) H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat Struct Mol Biol* 14:441–448
- Bustin M (2001a) Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem Sci* 26:152–153
- Bustin M (2001b) Chromatin unfolding and activation by HMGN chromosomal proteins. *Trends Biochem Sci* 26:431–437
- Butler PJG, Thomas JO (1980) Changes in chromatin folding in solution. *J Mol Biol* 140:505–529
- Cairney J, Booth IR, Higgins CF (1985) Osmoregulation of gene expression in *Salmonella typhimurium*: *proU* encodes an osmotically induced betaine transport system. *J Bacteriol* 164:1224–1232
- Caserta M, Agricola E, Churcher M, Hiriart E, Verdone L, Di Mauro E, Travers AA (2009) A translational signature for nucleosome positioning in vivo. *Nucleic Acids Res* 37:5309–5321
- Cheung KJ, Badarinarayana V, Selinger DW, Janse D, Church GM (2003) A microarray-based antibiotic screen identifies a regulatory role for supercoiling in the osmotic stress response of *Escherichia coli*. *Genome Res* 13:206–215
- Compton JL, Bellard M, Chambon P (1976) Biochemical evidence of variability in the DNA repeat length in the chromatin of higher eukaryotes. *Proc Natl Acad Sci USA* 73:4382–4386
- Costanzo G, Di Mauro E, Negra R, Pereira G, Hollenberg C (1995) Multiple overlapping positions of nucleosomes with single in vivo rotational setting in the *Hansenula polymorpha* RNA polymerase II MOX promoter. *J Biol Chem* 270:11091–11097
- Dame RT, Wyman C, Goosen N (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* 28:3504–3510
- Dorigo B, Schalch T, Bystricky K, Richmond TJ (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J Mol Biol* 327:85–96
- Dorigo B, Schalch T, Kulangara A, Duda S, Schroeder RR, Richmond TJ (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* 306:1571–1573
- Dorman CJ (1996) Flexible response: DNA supercoiling, transcription and bacterial adaptation to environmental stress. *Trends Microbiol* 4:214–216
- Doyen CM, Montel F, Gautier T, Menoni H, Claudet C, Delacour-Larose M, Angelov D, Hamiche A, Bednar J, Faivre-Moskalenko C, Bouvet P, Dimitrov S (2006) Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. *EMBO J* 25:4234–4244
- Drew HR, Travers AA (1985) DNA bending and its relation to nucleosome positioning. *J Mol Biol* 186:773–790
- Durrin LK, Mann RK, Kaye PS, Grunstein M (1992) Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* 65:1023–1031
- Eltsov M, Maclellan KM, Maeshima K, Frangakis AS, Dubochet J (2008) Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proc Natl Acad Sci USA* 105:19732–19737
- Felsenfeld G, McGhee JD (1986) Structure of the 30 nm chromatin fiber. *Cell* 44:375–379
- Finch JT, Klug A (1976) Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci USA* 73:1897–1901
- Finch JT, Lutter LC, Rhodes D, Brown RS, Rushton B, Levitt M, Klug A (1976) Structure of nucleosome core particles of chromatin. *Nature* 269:29–36
- Gautier T, Abbott DW, Molla A, Verdel A, Ausio J, Dimitrov S (2004) Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep* 5:715–720

- Gencheva M, Boa S, Fraser R, Simmen MW, Whitelaw CBA, Allan J (2006) In vitro and in vivo nucleosome positioning on the ovine β -lactoglobulin gene are related. *J Mol Biol* 361:216–230
- Gerchman SE, Ramakrishnan V (1987) Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. *Proc Natl Acad Sci USA* 84:7802–7806
- Germond JE, Bellard M, Oudet P, Chambon P (1976) Stability of nucleosomes in native and reconstituted chromatin. *Nucleic Acids Res* 3:3173–3192
- Ghirlando R, Felsenfeld G (2008) Hydrodynamic studies on defined heterochromatin fragments support a 30-nm fiber having six nucleosomes per turn. *J Mol Biol* 376:1417–1425
- Ghirlando R, Litt MD, Prioleau MN, Recillas-Targa F, Felsenfeld G (2004) Physical properties of a genomic condensed chromatin fragment. *J Mol Biol* 336:597–605
- Gralla JD, Vargas DR (2006) Potassium glutamate as a transcriptional inhibitor during bacterial osmoregulation. *EMBO J* 25:1515–1521
- Guo F, Adhya S (2007) Spiral structure of *Escherichia coli* HU α provides foundation for DNA supercoiling. *Proc Natl Acad Sci USA* 104:4309–4314
- Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, May G, Bremer E (1988) A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. *Cell* 52:569–584
- Hizume K, Yoshimura SH, Takeyasu K (2005) Linker histone H1 per se can induce three-dimensional folding of chromatin fiber. *Biochemistry* 44:12978–12989
- Hizume K, Nakai T, Araki S, Prieto E, Yoshikawa K, Takeyasu K (2009) Removal of histone tails from nucleosome dissects the physical mechanisms of salt-induced aggregation, linker histone H1-induced compaction, and 30-nm fiber formation of the nucleosome array. *Ultramicroscopy* epub March 19
- Hsieh CH, Griffith JD (1988) The terminus of SV40 DNA replication and transcription contains a sharp sequence-directed curve. *Cell* 52:535–544
- Izzo A, Kamieniarz K, Schneider R (2008) The histone H1 family: specific members, specific functions? *Biol Chem* 389:333–343
- Kan PY, Caterino TL, Hayes JJ (2009) The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Mol Cell Biol* 29:538–546
- Kar S, Edgar R, Adhya S (2005) Nucleoid remodeling by an altered HU protein: reorganization of the transcription program. *Proc Natl Acad Sci USA* 102:16397–16402
- Keller W, Müller U, Eicken I, Wendel I, Zentgraf H (1978) Biochemical and ultrastructural analysis of SV40 chromatin. *Cold Spring Harb Symp Quant Biol* 42:227–244
- Kepper N, Foethke D, Stehr R, Wedemann G, Rippe K (2008) Nucleosome geometry and internucleosomal interactions control the chromatin fiber conformation. *Biophys J* 95:3692–3705
- Kim Y, McLaughlin N, Lindstrom K, Tsukiyama T, Clark DJ (2006) Activation of *Saccharomyces cerevisiae* HIS3 results in Gcn4p-dependent, SWI/SNF-dependent mobilization of nucleosomes over the entire gene. *Mol Cell Biol* 26:8607–8622
- Kireeva N, Lakonishok M, Kireev I, Hirano T, Belmont AS (2004) Visualization of early chromosome condensation: a hierarchical folding, axial glue model of chromosome structure. *J Cell Biol* 166:775–785
- Kobryn K, Lavoie BD, Chaconas G (1999) Supercoiling-dependent site-specific binding of HU to naked Mu DNA. *J Mol Biol* 289:777–784
- Kornberg R (1981) The location of nucleosomes in chromatin: specific or statistical. *Nature* 292:279–280
- Kornberg RD, Stryer L (1988) Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Res* 16:6677–6689
- Krämer H, Niemöller M, Amouyal M, Revet B, von Wilcken-Bergmann B, Müller-Hill B (1987) *lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators. *EMBO J* 6:1481–1491
- Kruijthof M, Chien FT, Routh A, Logie C, Rhodes D, van Noort J (2009) Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nat Struct Mol Biol* 16:534–540

- Lang B, Blot N, Bouffartigues E, Buckle M, Geertz M, Gualerzi CO, Mavathur R, Muskhelishvili G, Pon CL, Rimsky S, Stella S, Babu MM, Travers A (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* 35:6330–6337
- Lazarus LR, Travers AA (1993) The *Escherichia coli* FIS protein is not required for the activation of *tyrT* transcription on entry into exponential growth. *EMBO J* 12:2483–2494
- Lee DH, Schleif RF (1989) In vivo DNA loops in *araCBAD*: size limits and helical repeat. *Proc Natl Acad Sci USA* 86:476–480
- Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C (2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* 39:1235–1244
- Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, Schumacher MA, Brennan RG, Lu P (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271:1247–1254
- Lu X, Simon MD, Chodaparambil JV, Hansen JC, Shokat KM, Luger K (2008) The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat Struct Mol Biol* 15:1122–1124
- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251–260
- Maurer S, Fritz J, Muskhelishvili G, Travers A (2006) RNA polymerase and an activator form discrete subcomplexes in a transcription initiation complex. *EMBO J* 25:3784–3790
- Maurer S, Fritz J, Muskhelishvili G (2009) A systematic in vitro study of nucleoprotein complexes formed by bacterial nucleoid associated proteins revealing novel types of DNA organization. *J Mol Biol* 387:1261–1276
- Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF (2008) A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res* 18:1073–1083
- McGhee JD, Nickol JM, Felsenfeld G, Rau DC (1983) Histone hyperacetylation has little effect on the higher order folding of chromatin. *Nucleic Acids Res* 11:4065–4075
- Muskhelishvili G, Travers A (2009) Intrinsic in vivo modulators: negative supercoiling and the constituents of the bacterial nucleoid. In: Buc H, Strick T (eds) *RNA polymerases as molecular motors*. RSC, Cambridge, pp 69–95
- Muskhelishvili G, Travers AA, Heumann H, Kahmann R (1995) FIS and RNA polymerase holoenzyme form a specific nucleoprotein complex at a stable RNA promoter. *EMBO J* 14:1446–1452
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313:236–238
- Nishioka K, Rice JC, Sarma K, Erdjument-Bromage H, Werner J, Wang Y, Chuikov S, Valenzuela P, Tempst P, Steward R, Lis JT, Allis CD, Reinberg D (2002) PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Mol Cell* 9:1201–1213
- Norton VG, Imai BS, Yau P, Bradbury EM (1989) Histone acetylation reduces nucleosome core particle linking number change. *Cell* 57:449–457
- Norton VG, Marvin KW, Yau P, Bradbury EM (1990) Nucleosome linking number change controlled by acetylation of histones H3 and H4. *J Biol Chem* 265:19848–19852
- Oberto J, Nabti S, Jooste V, Mignot H, Rouvière-Yaniv J (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS ONE* 4:e4367
- Olson WK, Gorin AA, Lu XJ, Hock LM, Zhurkin VB (1998) DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci USA* 95:11163–11168
- Paull TT, Johnson RC (1995) DNA looping by *Saccharomyces cerevisiae* high mobility group proteins NHP6A/B. Consequences for nucleoprotein complex assembly and chromatin condensation. *J Biol Chem* 270:8744–8754
- Pearson EC, Butler PJG, Thomas JO (1983) Higher-order structure of nucleosome oligomers from short-repeat chromatin. *EMBO J* 2:1367–1372

- Peck LJ, Wang JC (1981) Sequence dependence of the helical repeat of DNA in solution. *Nature* 292:375–378
- Pehrson JR (1995) Probing the conformation of nucleosome linker DNA in situ with pyrimidine dimer formation. *J Biol Chem* 270:22440–22444
- Pemberton IK, Muskhelishvili G, Travers AA, Buckle M (2002) FIS modulates the kinetics of successive interactions of RNA polymerase with the core and upstream regions of the *tyrT* promoter. *J Mol Biol* 318:651–663
- Piña B, Brüggemeier U, Beato M (1990a) Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. *Cell* 60:719–731
- Piña B, Truss M, Ohlenbusch H, Postma J, Beato M (1990b) DNA rotational positioning in a regulatory nucleosome is determined by base sequence. An algorithm to model the preferred superhelix. *Nucleic Acids Res* 18:6981–6987
- Pisano S, Pascucci E, Cacchione S, De Santis P, Savino M (2006) AFM imaging and theoretical modeling studies of sequence-dependent nucleosome positioning. *Biophys Chem* 124:81–89
- Pul U, Wurm R, Wagner R (2007) The role of LRP and H-NS in transcription regulation: involvement of synergism, allostery and macromolecular crowding. *J Mol Biol* 366:900–915
- Ragab A, Travers A (2003) HMG-D and histone H1 alter the local accessibility of nucleosomal DNA. *Nucleic Acids Res* 31:7083–7089
- Richard-Foy H, Hager GL (1987) Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *EMBO J* 6:2321–2328
- Robinson PJJ (2005) Reconstitution, folding and structure of the ‘30nm’ chromatin fibre. Ph.D. thesis, University of Cambridge, Cambridge, UK
- Robinson PJJ, Rhodes D (2006) Structure of the ‘30 nm’ chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* 16:336–343
- Robinson PJJ, Fairall L, Huynh VA, Rhodes D (2006) EM measurements define the dimensions of the “30-nm” chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci USA* 103:6506–6511
- Robinson PJJ, An W, Routh A, Martino F, Chapman L, Roeder RG, Rhodes D (2008) 30 nm chromatin fibre decompaction requires both H4–K16 acetylation and linker histone eviction. *J Mol Biol* 381:816–825
- Routh A, Sandin S, Rhodes D (2008) Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc Natl Acad Sci USA* 105:8872–8877
- Rouvière-Yaniv J, Yaniv M, Germond JE (1979) *E. coli* DNA binding protein HU forms nucleosomelike structure with circular double-stranded DNA. *Cell* 17:265–274
- Rydberg B, Holley WR, Mian IS, Chatterjee A (1998) Chromatin conformation in living cells: support for a zig-zag model of the 30 nm chromatin fiber. *J Mol Biol* 284:71–84
- Satchwell SC, Drew HR, Travers AA (1986) Sequence periodicities in chicken nucleosome core DNA. *J Mol Biol* 191:659–675
- Saviola B, Seabold PR, Schleif RF (1998) DNA bending by AraC: a negative mutant. *J Bacteriol* 180:4227–4232
- Schalch T, Duda S, Sargent DF, Richmond TJ (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436:138–141
- Schneider R, Lurz R, Lüder G, Tolksdorf C, Travers A, Muskhelishvili G (2001) An architectural role of the *Escherichia coli* protein FIS in organising DNA. *Nucl Acids Res* 29:5107–5114
- Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, Wang JP, Widom J (2006) A genomic code for nucleosome positioning. *Nature* 442:772–778
- Shen CH, Clark DJ (2001) DNA sequence plays a major role in determining nucleosome positions in yeast CUP1 chromatin. *J Biol Chem* 276:35209–35216
- Shimizu M, Roth SY, Szent-Gyorgyi C, Simpson RT (1991) Nucleosomes are positioned with base pair precision adjacent to the alpha 2 operator in *Saccharomyces cerevisiae*. *EMBO J* 10:3033–3041
- Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL (2006) Histone H4–K16 acetylation controls chromatin structure and protein interactions. *Science* 311:844–847
- Shrader TE, Crothers DM (1989) Artificial nucleosome positioning sequences. *Proc Natl Acad Sci USA* 86:7418–7422

- Staynov DZ, Proykova YG (2008) Topological constraints on the possible structures of the 30 nm chromatin fibre. *Chromosoma* 117:67–76
- Swinger KK, Lemberg KM, Zhang Y, Rice PA (2003) Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J* 22:3749–3760
- Tanaka S, Zatchej M, Thoma F (1992) Artificial nucleosome positioning sequences tested in yeast minichromosomes: a strong rotational setting is not sufficient to position nucleosomes in vivo. *EMBO J* 11:1187–1193
- Thoma F, Koller T, Klug A (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol* 83:403–427
- Thomas JO (1999) Histone H1: location and role. *Curr Opin Cell Biol* 11:312–317
- Thomas JO, Butler PJG (1980) Size-dependence of a stable higher-order structure of chromatin. *J Mol Biol* 144:89–93
- Thomas JO, Kornberg RD (1975) An octamer of histones in chromatin and free in solution. *Proc Natl Acad Sci USA* 72:2626–2630
- Thomas JO, Thompson RJ (1977) Variation in chromatin structure in two cell types from the same tissue: a short DNA repeat length in cerebral cortex neurons. *Cell* 10:633–640
- Thomas JO, Travers AA (2001) HMG1 and 2, and related ‘architectural’ DNA-binding proteins. *Trends Biochem Sci* 26:167–174
- Tóth KF, Knoch TA, Wachsmuth M, Frank-Stöhr M, Stöhr M, Bacher CP, Müller G, Rippe K (2004) Trichostatin A-induced histone acetylation causes decondensation of interphase chromatin. *J Cell Sci* 117:4277–4287
- Travers A (2007) DNA bending and nucleosome positioning. *Trends Biochem Sci* 12:108–112
- Travers AA (2005) Gene regulation by HMGA and HMGB chromosomal proteins and related architectural DNA-binding proteins. In: Ohyama T (ed) *DNA conformation in transcription*. Landes Bioscience, Georgetown, pp 147–158
- Travers A (2009) In: Benham CJ, Harvey S, Olson W, Summers DW (eds) *IMA volume 150: mathematics of DNA structure, function, and interactions*. Springer Science and Business Media, New York, pp 321–330
- Travers A, Klug A (1987) The bending of DNA in nucleosomes and its wider implications. *Philos Trans R Soc Lond B Biol Sci* 317:5375–5361
- Travers A, Muskhelishvili G (2005a) DNA supercoiling – a global transcriptional regulator for enterobacterial growth? *Nature Rev Microbiol* 3:157–169
- Travers A, Muskhelishvili G (2005b) Bacterial chromatin. *Curr Opin Genet Dev* 15:507–514
- Travers A, Muskhelishvili G (2007) A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep* 8:147–151
- Travers AA, Ner SS, Churchill MEA (1994) DNA chaperones: a solution to a persistence problem? *Cell* 77:167–169
- Travers A, Caserta M, Churcher M, Hiriart E, Di Mauro E (2009) Nucleosome positioning – what do we know? *Mol. BioSystems* doi:10.1039/B907227F
- Tse C, Sera T, Wolffe AP, Hansen JC (1998) Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol Cell Biol* 18:4629–4638
- Tupper AE, Owen-Hughes TA, Ussery DW, Santos DS, Ferguson DJ, Sidebotham JM, Hinton JC, Higgins CF (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J* 13:258–268
- Urata Y, Parmelee SJ, Agard DA, Sedat JW (1995) A three-dimensional structural dissection of *Drosophila* polytene chromosomes. *J Cell Biol* 131:279–295
- Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A, Johnson SM (2008) A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res* 18:1051–1063
- Virstedt J, Berge T, Henderson RM, Waring MJ, Travers AA (2004) The influence of DNA stiffness upon nucleosome formation. *J Struct Biol* 148:66–85

- Wallrath LL, Lu Q, Granok H, Elgin SCR (1994) Architectural variations of inducible eukaryotic promoters: preset and remodeling chromatin structures. *BioEssays* 16:165–170
- Wang X, Hayes JJ (2008) Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure. *Mol Cell Biol* 28:227–236
- Weiss K, Simpson RT (1997) Cell type-specific chromatin organization of the region that governs directionality of yeast mating type switching. *EMBO J* 16:4352–4360
- Widom J (1986) Physicochemical studies of the folding of the 100 Å nucleosome filament into the 300 Å filament Cation dependence. *J Mol Biol* 190:411–424
- Widom J (2001) Role of DNA sequence in nucleosome stability and dynamics. *Q Rev Biophys* 34:269–324
- Widom J, Klug A (1985) Structure of the 300Å chromatin filament: X-ray diffraction from oriented samples. *Cell* 43:207–213
- Williams SP, Athey BD, Muglia LJ, Schappe RS, Gough AH, Langmore JP (1986) Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys J* 49:233–248
- Wolf SG, Frenkiel D, Arad T, Finkel SE, Kolter R, Minsky A (1999) DNA protection by stress-induced biocrystallization. *Nature* 400:83–85
- Wolffe AP, Drew HR (1989) Initiation of transcription on nucleosomal templates. *Proc Natl Acad Sci USA* 86:9817–9821
- Wong H, Victor JM, Mozziconacci J (2007) An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length. *PLoS ONE* 2:e877
- Woodcock CL, Frado LL, Rattner JB (1984) The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol* 99:42–52
- Worcel A, Strogatz S, Riley D (1981) Structure of chromatin and the linking number of DNA. *Proc Natl Acad Sci USA* 78:1461–1465
- Wu C, Bassett A, Travers A (2007) A variable topology for the 30-nm chromatin fibre. *EMBO Rep* 8:1129–1134
- Zhou J, Fan JY, Rangasamy D, Tremethick DJ (2007) The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat Struct Mol Biol* 14:1070–1076
- Zivanovic Y, Goulet I, Revet B, Le Bret M, Prunell A (1988) Chromatin reconstitution on small DNA rings II. DNA supercoiling on the nucleosome. *J Mol Biol* 200:267–290

Part III
Regulation by Nucleoid-Associated
Proteins

Chapter 12

Bacterial Chromatin and Gene Regulation

Charles J. Dorman

Abstract Bacterial chromatin is composed of the genetic material of the cell with its associated proteins and RNA molecules. The proteins provide structural and regulatory functions and frequently both functions can be found in individual polypeptides. Nucleoid-associated proteins form a very important group that combines genome architectural and gene regulatory functions. They are highly versatile and make important contributions to the control of transcription at a global level. In combination with DNA structural features such as supercoiling, nucleoid-associated proteins help to set the scene against which more focused aspects of gene control must operate. The combination of DNA structure, nucleoid-associated proteins and conventional transcription regulators provides the basis for an integrated gene control network in bacteria that is both environmentally responsive and highly adaptable.

Keywords Nucleoid-associated proteins • DNA supercoiling • transcription • gene regulation • H-NS • FIS • HU • IHF • DPS • CRP • LRP

Studies of bacterial chromatin highlight for us the interwoven nature of nucleoid structure and gene regulation. Most of the elements that contribute to nucleoid structure have the potential to influence transcription, either directly or indirectly. Furthermore, the major transactions that take place within DNA, including transcription, influence nucleoid structure. Molecular crowding, DNA supercoiling and the association of DNA with architectural proteins all make essential contributions to the organization of the nucleoid and they have also been implicated in modulating gene expression. This duality points to the integrated nature of the cell and it makes studies of the nucleoid a very attractive research topic for those interested in understanding how a bacterium functions as a biological unit.

Until relatively recently, investigations of nucleoid structure and gene regulation were conducted at different levels within the organization of the cell. Studies of the

C.J. Dorman (✉)

Department of Microbiology, Moyne Institute of Preventive Medicine,
School of Genetics and Microbiology, Trinity College Dublin, Dublin 2, Ireland
e-mail: cjdorman@tcd.ie

nucleoid typically relied (and still rely) on research techniques that consider the entire chromosome, with or without its attendant proteins, *in vivo* or *in vitro*. On the other hand, studies of gene regulation have often involved detailed dissection of the molecular events that influence the activity of a single promoter or a small number of promoters. The emergence of methods for surveying promoter activity throughout the genome has ushered in a new era of global gene regulation research. Transcriptomic, proteomic and metabolomic methodologies, allied with powerful *in silico* bioinformatic capabilities are providing important insights into the relationships between gene regulatory elements and the promoters that they control (Balleza et al. 2009). They also offer the possibility of conducting the work in situations that are relevant to the everyday lives of the bacteria that we study (Verberkmoes et al. 2009).

Nucleoid-associated proteins (NAPs) are a major topic of this book and they are known to influence the expression of many bacterial genes (Ali Azam et al. 1999; Dorman and Deighan 2003; Drlica and Rouvière-Yaniv 1987). These influences are felt chiefly, but not exclusively, at the level of transcription, as one might predict for a class of proteins whose primary common activity is DNA binding. Their roles in gene expression emerged in many cases from genetic studies that sought to identify *trans*-acting regulators of particular genes or important components of the bacterial response to certain signals or stresses. One should not overlook the very important role that has been played by genetic studies of the life cycles of bacteriophage (especially lambda) and the biology of plasmids in the elucidation of the contributions made by NAPs to the regulatory networks of bacteria (Oppenheim et al. 2005). It is also useful to recall that much of this pioneering work was done before the complete genome sequence of even one bacterium had been published. Despite the limitations imposed by the dearth of whole genome sequences, some investigators pursued global regulatory studies using genetic tools that allowed reporter genes to be delivered throughout the chromosome at random by transposons or phage and these studies generated useful data about the nature of global networks (Errington 1986; Jovanovich and Lebowitz 1987).

The early history of gene regulation in bacteria was dominated by a small number of model systems such as *lac* or *trp* that led to very important understandings of transcriptional and posttranscriptional control through attenuation, repression, the nature of allostery and other essential concepts that are still pillars of cell biology today (Brock 1990). These models were the first to appear in the textbooks and in undergraduate (and postgraduate) courses on molecular genetics and have created a strong impression of what ‘conventional’ gene regulation looks like. At least two generations of investigators have imbibed this conventional view and it has served us very well indeed. This tradition has provided a solid foundation for the new global thinking about regulatory networks, but it did not predict the emerging realization that architectural elements from the nucleoid play such a central role in controlling major gene expression programmes.

Leading experts have contributed the chapters in this book and they provide up-to-date surveys of many aspects of bacterial chromatin. Their writing illustrates very elegantly the integrated nature of the structural and regulatory aspects of bacterial chromatin. Each chapter has something to say about each aspect, although several are more regulatory in nature while others take a more structural approach.

Even a cursory glance at the contents of the book will show how advanced the NAP field has become in *Escherichia coli* and its relatives and how intimately these proteins are associated with gene regulatory processes. The reader will find chapters dedicated to the NAPs Dps, FIS, H-NS, HU, IHF, LRP, and will discover that these same proteins appear in more than one chapter due to the pervasive nature of their influence on bacterial biology. There is also a discussion of CRP, a protein discovered very early in the history of molecular biology that some may not have considered before as a NAP. As one of the first global regulators to be studied in detail CRP is readily associated with catabolite repression and the control of operations involved in the uptake and utilization of carbohydrates (Fic et al. 2006). Recently, transcriptomic work and experiments with chromatin-immunoprecipitation-on-chip, led by Steve Busby, have revealed to us just how truly widespread is the influence of CRP in the cell (Grainger et al. 2005). This protein stands at the interface between our views of conventional gene regulation and global control by NAPs, and calls into question our working definition of NAPs. Biologists are comfortable with classification systems and require clear rules that permit unambiguous assignments, whether dealing with molecules or whole organisms. One of the messages of our book is that facile assignments are unlikely to be available for gene regulatory proteins in bacteria. Workers in the field may be confident that they know a NAP when they see one, but precise definitions are elusive.

Conventional transcription factors are typically expressed in relatively small amounts and in many cases their activities can be modulated by ligand binding or covalent modification. Might these features mark points of distinction between transcription factors and NAPs? Proteins like FIS, HU, IHF, and H-NS (to name but four NAPs) are abundant and are not known to undergo allosteric regulation (Atlung and Ingmer 1997; Dorman 2009; Grainger and Busby 2008; Swinger and Rice 2004; Yokoyama et al. 2006). In contrast, CRP and LRP both bind signalling molecules – cAMP in the case of CRP and leucine in the case of LRP – that alter their biological activity. In the case of CRP, DNA binding is contingent on binding of cAMP by the protein; in the case of LRP, binding of leucine can have positive, negative or no influence on the ability of this protein to affect transcription – it all depends on the LRP-dependent promoter in question (Fic et al. 2006; Yokoyama et al. 2006). Attempts at definitions based on the specificity of DNA binding do not help much either: CRP and IHF have clearly defined preferences for binding to particular DNA sequences but this is not the case with many other NAPs, such as FIS, H-NS and HU.

Perhaps it is more productive to consider a role for gene regulators in nucleoid structuring as a qualification for membership of the NAP family. Various attempts have been made to identify proteins that form the boundaries that close the independent loops that arise stochastically in the folded chromosomes of bacteria. NAPs like H-NS and FIS make very attractive candidates, not least because they possess a DNA bridging activity that lends itself to this role (Dame et al. 2005; Luijsterburg et al. 2008). Their involvement has support from genetic studies as well as strong circumstantial evidence from chromatin immunoprecipitation work. However, there is also genetic evidence in favour of a domain boundary role for proteins that lack DNA binding activity such as DksA, phosphoglucosyltransferase and transketolase (Hardy and Cozzarelli 2005). In summary, attempts at narrowly

defined definitions do not serve us very well and it may be more productive to consider that model bacteria like *E. coli* possess a large population of DNA binding proteins that possess a spectrum of activities that encompass structural and regulatory roles. Among these are some that have truly global effects on gene expression. What is the biological value of these global regulators?

Some NAPs may set limits within which the transcription programme of the cell has to operate. They may do this directly or they may act in association with aspects of DNA topology such as supercoiling (Dorman 2006). FIS, HU and H-NS are three examples. The widespread and largely negative influence of H-NS on the transcriptome is reviewed comprehensively in the chapter by William Navarre. This negative influence is a fact of life for bacteria that express H-NS or an analogous protein and ways must be found to displace it or to allow the transcription machinery to work around it if H-NS-repressed genes are to be expressed. The list of antagonists that lift H-NS-imposed repression is long and growing and some NAPs are included (Stoebe et al. 2008). Both FIS and HU have the ability to remodel DNA that has been bridged by H-NS and both of these proteins make important contributions to transcription regulation as discussed by Rolf Wagner, Georgi Muskhelishvili, Andrew Travers (FIS) and by Sankar Adhya and colleagues (HU) (Dame and Goosen 2002; Maurer et al. 2009). Other NAPs with the potential to do the same include LRP (Stacey Peterson and Norbert Reich) and IHF (Victor de Lorenzo and colleagues). The Dps protein seems to be in a class by itself where gene regulation is concerned (see the chapter by Hanne Ingmer). Its gene is subject to regulation by the NAPs H-NS, FIS and IHF but it does not seem to be an important regulator of transcription (Altuvia et al. 1994; Grainger et al. 2008). This is both surprising and a very interesting observation for a very abundant protein with general DNA binding activity (Frenkiel-Krispin and Minsky 2006). Dps illustrates once again the difficulty of providing an all-encompassing description of the global regulators of transcription in bacteria like *E. coli*; a protein that one might predict to be an excellent candidate for the role of transcription repressor under those conditions when it is abundant (oxidative stress or stationary phase) does not appear to have the anticipated properties.

The involvement of nucleoid structuring elements in gene regulation raises interesting questions about the evolution of regulation and about the potential of bacteria to evolve new regulatory regimes in the future. Negative regulation by the H-NS protein may be a by-product of its structural role and the fortuitous presence of the features of DNA that attract it (high A + T content and a nearby region of curved DNA) at bacterial promoters (Dorman 2007). The DNA bridging activity of H-NS can be opposed by alternative interactions with DNA (wrapping and/or bending) that are exhibited by other DNA binding proteins (Luijsterburg et al. 2008; Stoebe et al. 2008). This antagonism provides the basis for a primitive genetic switch and the operation of the switch can be modulated by influencing the structure of the DNA, and/or the activities of the structural/regulatory proteins and/or their supply. We see examples of all of these strategies at work in modern model bacterial cells and the reader will see them described in the pages of this book. Their very simplicity lends itself to further development under the pressures imposed by natural selection. Understanding how bacteria evolve new gene regulatory regimes within the con-

straints imposed by the structural needs of the nucleoid will keep the topic of bacterial chromatin at the forefront of molecular and cell biology for the foreseeable future.

Acknowledgements Research in the author's laboratory is supported by grants from Science Foundation Ireland and the Wellcome Trust.

References

- Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181:6361–6370
- Atlung T, Ingmer H (1997) H-NS: a modulator of environmentally regulated gene expression. *Mol Microbiol* 24:7–17
- Balleza E, López-Borrorquez LN, Martínez-Antonio A, Resendis-Antonio O, Lozada-Chávez I, Balderas-Martínez YI, Encarnación S, Collado-Vides J (2009) Regulation by transcription factors in bacteria: beyond description. *FEMS Microbiol Rev* 33:133–151
- Brock TD (1990) The emergence of bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- Dame RT, Goosen N (2002) HU: promoting or counteracting DNA compaction? *FEBS Lett* 529:151–156
- Dame RT, Luijsterburg MS, Krin E, Bertin PN, Wagner R, Wuite GJ (2005) DNA bridging: a property shared among H-NS-like proteins. *J Bacteriol* 187:1845–1858
- Dorman CJ (2006) DNA supercoiling and bacterial gene expression. *Sci Prog* 89:151–166
- Dorman CJ (2007) H-NS, the genome sentinel. *Nat Rev Microbiol* 5:157–161
- Dorman CJ (2009) Nucleoid-associated proteins and bacterial physiology. *Adv Appl Microbiol* 67:47–64
- Dorman CJ, Deighan P (2003) Regulation of gene expression by histone-like proteins in bacteria. *Curr Opin Genet* 13:179–184
- Drlica K, Rouvière-Yaniv J (1987) Histone-like proteins of bacteria. *Microbiol Rev* 51:301–319
- Errington J (1986) A general method for fusion of the *Escherichia coli* *lacZ* gene to chromosomal genes in *Bacillus subtilis*. *J Gen Microbiol* 132:2953–2966
- Fic E, Polit A, Wasylewski Z (2006) Kinetic and structural studies of the allosteric conformational changes induced by binding of cAMP to the cAMP receptor protein from *Escherichia coli*. *Biochemistry* 45:373–380
- Frenkiel-Krispin D, Minsky A (2006) Nucleoid organization and the maintenance of DNA integrity in *E. coli*, *B. subtilis* and *D. radiodurans*. *J Struct Biol* 156:311–319
- Grainger DC, Busby SJW (2008) Global regulators of transcription in *Escherichia coli*: mechanisms of action and methods for study. *Adv Appl Microbiol* 65:93–113
- Grainger DC, Hurd D, Harrison M, Holdstock J, Busby SJW (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci USA* 102:17693–17698
- Grainger DC, Goldberg MD, Lee DJ, Busby SJW (2008) Selective repression by Fis and H-NS at the *Escherichia coli* *dps* promoter. *Mol Microbiol* 68:1366–1377
- Hardy CD, Cozzarelli NR (2005) A genetic selection for supercoiling mutants of *Escherichia coli* reveals proteins implicated in chromosome structure. *Mol Microbiol* 57:1636–1652
- Jovanovich SB, Lebowitz J (1987) Estimation of the effect of coumermycin A1 on *Salmonella typhimurium* promoters by using random operon fusions. *J Bacteriol* 169:4431–4435
- Luijsterburg MS, White MF, van Driel R, Dame RT (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol* 43:393–418
- Maurer S, Fritz J, Muskhelishvili G (2009) A systematic in vitro study of nucleoprotein complexes formed by bacterial nucleoid-associated proteins revealing novel types of DNA organization. *J Mol Biol* 387:1261–1276

- Oppenheim AB, Kobiler O, Stavans J, Court DL, Adhya S (2005) Switches in bacteriophage lambda development. *Annu Rev Genet* 39:409–429
- Stoebel DM, Free A, Dorman CJ (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* 154:2533–2545
- Swinger KK, Rice PA (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol* 14:28–35
- Verberkmoes NC, Denef VJ, Hettich RL, Banfield JF (2009) Systems biology: functional analysis of natural microbial consortia using community proteomics. *Nat Rev Microbiol* 7:196–205
- Yokoyama K, Ishijima SA, Clowney L, Koike H, Aramaki H, Tanaka C, Makino K, Suzuki M (2006) Feast/famine regulatory proteins (FFRPs): *Escherichia coli* Lrp, AsnC and related archaeal transcription factors. *FEMS Microbiol Rev* 30:89–108

Chapter 13

H-NS as a Defence System

William Wiley Navarre

Abstract The genomes of free-living bacteria are not static but, quite to the contrary, are highly dynamic entities over evolutionary time. Bacteria evolve rapidly by constantly swapping genetic material through lateral (horizontal) gene transfer. Even potentially useful foreign DNA can interfere with pre-existing regulatory networks and disrupt nucleoid structure and chromatin organization. The nucleoid associated protein H-NS plays a major role in binding and silencing expression from genes acquired by enteric bacteria from foreign sources by specifically targeting those sequences with significantly lower GC-content than the rest of the genome. As a result of this property H-NS regulates a large majority of virulence-associated genes in enteric bacteria including *Salmonella*, *E. coli*, and *Yersinia*. This chapter will focus on what is known about H-NS-mediated silencing, anti-silencing by H-NS antagonists, H-NS phylogeny, the accessory molecules Hha and YdgT, and the role H-NS plays in protecting the bacterial cell from the negative consequences of genetic exchange.

Keywords H-NS • Xenogeneic silencing • Lateral gene transfer • Horizontal gene transfer • Nucleoid structuring protein • DNA topology

13.1 Introduction

The sequencing of thousands of bacterial genomes over the past decade has led to a revolution in our understanding of how these unicellular organisms evolve. We have learned that bacterial chromosomes are not static but, quite to the contrary, can display large variations in gene content even between strains that are closely related. This striking amount of genetic variability is due in large part to the fact that the genomes of free-living bacteria frequently undergo genetic exchange via

W.W. Navarre (✉)

Department of Molecular Genetics, University of Toronto, 1 King's College
Circle – Room 4379, Toronto, Ontario, M5S 1A8, Canada
e-mail: william.navarre@utoronto.ca

lateral (or horizontal) gene transfer (LGT), picking up new genetic material through processes such as conjugation and transduction while rapidly losing genes that do not confer a selective advantage (Ochman and Davalos 2006; van Passel et al. 2008). The role of LGT in bacterial evolution has been primarily studied in the context of pathogens as a large majority of virulence-associated genes in most pathogens were acquired by LGT. It is clear, however, that genetic exchange is also a commonplace phenomenon in non-pathogenic bacteria and evidence of LGT has been observed in the genomes of every free-living bacterial species sequenced thus far.

Any *E. coli* strain can differ from any *Salmonella* strain by as much as a quarter of its genetic content despite the fact that the genomes are largely conserved in their gene order (i.e. are syntenic) and only diverged approximately 100 MYr ago (Groisman and Ochman 1997; Lawrence and Ochman 1998; McClelland et al. 2001; van Passel et al. 2008). Individual isolates of other bacterial species, such as members of the high GC Gram-positive *Frankia*, can vary in their genetic content from one another by nearly 50% (Normand et al. 2007). All of this begs the question of how bacterial cells can organize, structure, and regulate their genome in the background of extensive and rapid gene acquisition by LGT and reduction through gene loss.

Genes acquired by LGT can be associated with either intact or degraded phage elements or are located adjacent to tRNA loci, but many other genes appear to have inserted into the genome with no apparent clue as to the underlying mechanism of their transfer. Xenogeneic sequences frequently exhibit variant base composition and codon usage from that of their host genome; in a majority of cases, genes acquired by LGT are relatively rich in adenine and thymine (i.e. are “AT-rich”) (Daubin et al. 2003a; Lawrence and Ochman 1997).

This chapter will focus on the bacterial nucleoid protein H-NS and the recent evidence that suggests it plays a primary role in targeting and silencing those genes obtained via LGT with AT contents significantly higher than the genome average. This chapter will explore specific properties that make H-NS ideally suited to serve as a silencer of gene expression and the mechanisms by which H-NS-silenced genes are activated. Despite recent progress and almost three decades of intense study, a complete understanding of H-NS and its function has been elusive and virtually every aspect of H-NS structure and function remains controversial to one degree or another.

13.2 H-NS

H-NS was originally identified as a small heat-stable factor that could stimulate *E. coli* RNA polymerase (RNAP)-directed transcription in vitro from phage templates when present at low concentrations but would inhibit transcription at high concentrations (Cukier-Kahn et al. 1972; Jacquet et al. 1971). H-NS was rediscovered in 1977 and yet again in 1981 during biochemical screens for bacterial “histone-like” proteins that could be isolated under conditions that were successful in isolating eukaryotic histones (Bakaev 1981; Varshavsky et al. 1977). It was named H-NS

(heat-stable nucleoid structuring protein) during preliminary structural and functional studies of bacterial “histone-like” proteins (Falconi et al. 1988; Gualerzi et al. 1986; Lammi et al. 1984; Paci et al. 1986) and was concurrently identified in several independent studies as responsible for regulating a large number of disparate biological processes. As such H-NS and its corresponding gene, *hns*, have been at some point referred to as H1 (Cukier-Kahn et al. 1972), H1a (Spassky et al. 1984), 16K (Laine et al. 1984), B1 (Varshavsky et al. 1977), *bglY* (Defez and De Felice 1981), *osmZ* (May et al. 1990), *drdX* (Goransson et al. 1990), *virR* (Hromockyj et al. 1992), *cur* (Diderichsen 1980) and *pilG* (Spears et al. 1986). Any confusion regarding the name of this gene and its product was resolved by the early 1990s and *hns*/H-NS, with a few unfortunate exceptions, has been used almost exclusively since that time.

H-NS has frequently been called one of the “histone-like proteins of bacteria”. This statement carries not only the misconception that H-NS is functionally equivalent to a histone, but also that it is a molecule found in most bacteria. Although several comparisons have been made between H-NS and eukaryotic histones, and despite the fact that they share some gross overall characteristics like the ability to compact and increase the thermal stability of DNA (Friedrich et al. 1988; Lammi et al. 1984; Spassky et al. 1984), it is important to note that H-NS bears no primary sequence homology to any histone subunit nor is H-NS similar to histones in the manner by which it interacts with DNA.

Analysis of available genome sequences has revealed that members of the family of H-NS-like molecules are incredibly diverse in their primary sequence. Many are so divergent that they cannot be easily aligned with H-NS for much of their sequence nor can they be identified through simple BLAST searches (Tendeng and Bertin 2003; Tendeng et al. 2003b). The distribution of H-NS-like molecules is limited to subsets of the alpha-, beta-, and gamma-proteobacteria. Furthermore the distribution of these proteins is quite odd with many H-NS like molecules encoded on genomic islands or plasmids that have been acquired via LGT. *E. coli* can encode between two and four H-NS like molecules while species like *Burkholderia* can encode as many as eighteen. The implications of these observations are discussed further in section 11 regarding H-NS phylogeny.

13.3 Structure of H-NS

H-NS is a small (15.5 kDa, 137 amino acids) neutral DNA-binding protein that is one of the most abundant proteins in *E. coli* at approximately 20,000 copies per cell (Falconi et al. 1988; Lammi et al. 1984). A complete, high-resolution structure of the H-NS protein has proven elusive but data from several studies support a model where H-NS consists of an N-terminal domain and flexible linker that are involved in multimerization and higher order structuring and a C-terminal DNA-binding domain (Rimsky 2004; Shindo et al. 1999; Smyth et al. 2000; Ueguchi et al. 1996, 1997). The N-terminal domain is contained within residues 1–64 while the C-terminal domain extends approximately from residues 80 to 136. Structures of the N-terminal (Bloch et al. 2003;

Esposito et al. 2002; Renzoni et al. 2001) and C-terminal domains (Shindo et al. 1995, 1999) from the *E. coli* or *Salmonella* H-NS molecules have been resolved by NMR analysis, but no high-resolution molecular structure exists of full-length H-NS in complex with target DNA. The salient features of each domain are explained in greater detail in the following sections with a central focus on how they contribute to the ability of H-NS to act as a transcriptional silencer.

13.3.1 Dimerization and Higher-Order Complexes – A Role for the N-Terminus

Several studies have demonstrated that the N-terminal domain of H-NS plays a role in H-NS multimerization but despite considerable effort by several labs we still have a very unclear picture of how H-NS self-associates. Interpretations regarding the role of specific regions and even the degree of multimerization vary widely depending on the approach used. An early study by Falconi et al. found that in solution H-NS is predominately dimeric above concentrations of 10 μ M and that at even higher concentrations a significant amount of H-NS can be found in trimers and tetramers (Falconi et al. 1988). In this study the presence of DNA did not significantly affect the multimerization state of H-NS. Another study by Ueguchi et al. observed dimers and tetramers in a concentration dependent manner as determined by the elution profile through a gel-filtration matrix (Ueguchi et al. 1996). A similar approach on the isolated N-terminal domain (residues 1–64) using both gel filtration and analytical ultracentrifugation found that the N-terminal domain exists primarily as a trimer in solution and that the full-length protein can form $\approx$ 20mers in solution at sufficiently high concentration (0.34 mM) (Smyth et al. 2000), but this finding is disputed by a similar study of the same domain that found the 1–64 region forms a dimer during analytical ultracentrifugation (Esposito et al. 2002). Yet another study in which large-zone gel permeation was employed found that dimers exist only transiently and that the majority of the protein could be found in the tetrameric form at higher protein concentrations (Ceschini et al. 2000). These examples illustrate the difficulty that has plagued structure/function studies on H-NS using traditional biochemical approaches.

High-resolution solution structures of the N-terminal domain from *E. coli* (residues 1 to 46) (Bloch et al. 2003) and *Salmonella enterica* Serovar. Typhimurium (residues 1–64) (Esposito et al. 2002; Renzoni et al. 2001) have been solved by multi-dimensional NMR. Another N-terminal domain structure (residues 2–49) from the H-NS-like VicH protein of *Vibrio cholerae* was solved through crystallography (Cerdan et al. 2003). In agreement with earlier studies of this domain by circular dichroism, these studies each found that the N-terminal domain is alpha helical, composed of two short alpha helices followed by a longer one. The structures also agree with other findings that the N-terminal domain forms a homodimer. The agreements appear to end there, however, as the *Vibrio* and *E. coli* structures have the dimer arranged in an “antiparallel” handshake arrangement (Fig. 13.1) while the *Salmonella* structure finds the two dimers are arranged in parallel – an arrangement

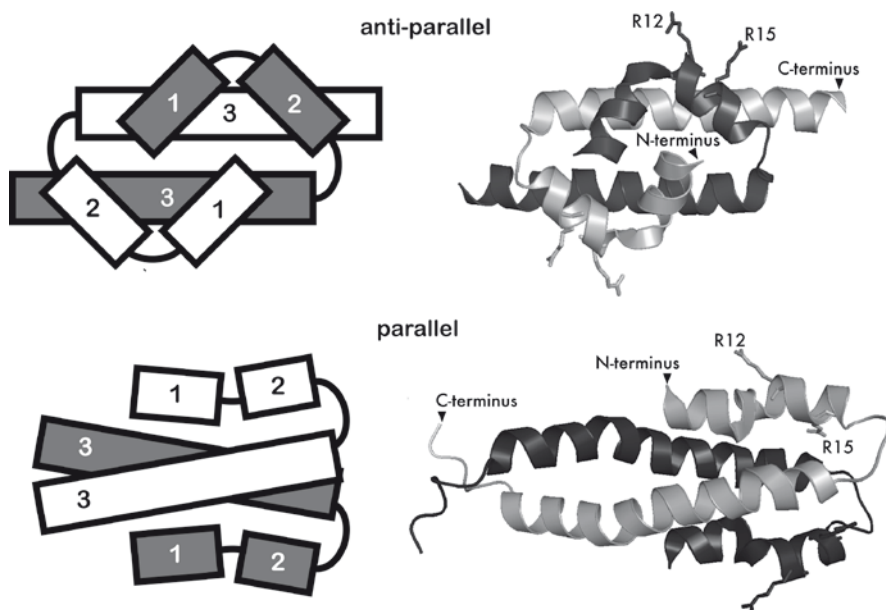


Fig. 13.1 Structure and possible configurations of the H-NS N-terminal dimerization domain. Solution structures that have been generated for the N-terminal domain of the *Escherichia coli* H-NS (residues 1–46) (Bloch et al. 2003) and *Salmonella enterica* Sv. Typhimurium H-NS (residues 1–64) (Esposito et al. 2002) molecules have come to strikingly different conclusions regarding their tertiary arrangement. The shorter *E. coli* fragment was found to exist as a dimer with the subunits arranged in an anti-parallel “handshake fold” (top panel). The longer *S. Typhimurium* fragment was determined to adopt a “parallel” configuration as depicted in the right panel. Individual subunits in each dimer are labeled light or dark grey. Diagrams of the three α -helices (labeled 1, 2 and 3) and their respective arrangements in each structural model are shown to the left of each ribbon structure. The side-chain of arginine 12 within the highly conserved motif (NNIRTL) found in the H-NS molecules from the *enterobacteriaceae* necessary for interactions with Hha is shown (Garcia et al. 2006; Madrid et al. 2007b). The side chain of arginine 15, that is critical for H-NS dimerization is also indicated (Garcia et al. 2006)

also inferentially supported by measurements derived from another NMR study and single molecule manipulations of H-NS complexed to DNA (Dame et al. 2006; Garcia et al. 2006). Both proposed arrangements share certain central and important features. The primary interaction between monomers occurs through a hydrophobic coiled-coil interaction using residues contained primarily in helix 3 and several charged residues are also predicted to interact with one another via salt bridges. The two proposed dimer arrangements differ almost completely in the pairings between critical interacting residues. For example, in the parallel structure the arginine at position 15 is predicted to pair with glutamates at positions 24 and 27, while in the anti-parallel structure this same arginine is paired with the glutamate at position 39. Both structures predict that the surface of the dimerization domain is negatively charged, suggesting that it is unlikely to interact directly with DNA.

13.3.2 DNA Binding – A Role for the C-Terminal Domain of H-NS

The structure of the C-terminal domain (residues 91–137) of H-NS was resolved through NMR analysis (Shindo et al. 1995). This 47 residue fragment was isolated by limited trypsin proteolysis and was shown to retain partial DNA binding activity, albeit with approximately 3 orders of magnitude less affinity than full-length H-NS. The structure consists of a two-stranded beta-sheet, an alpha helix, and a 3_{10} helix, each separated by small loops. A follow up study used NOE NMR to measure the interaction of a C-terminal H-NS fragment containing residues 60–137 with a 14 basepair oligonucleotide duplex. The H-NS residues that interact with DNA were localized to two separate areas: a disordered loop preceding the beta-sheet region (residues A80 to K96) and the loop between the beta-sheet and the alpha-helix regions (T110 to A117) (Shindo et al. 1999). These loops are adjacent in the folded structure and contain several positively charged residues that likely facilitate their interaction with DNA. The DNA interaction domain as defined in the NOE NMR study is consistent with an extensive mutagenesis study that identified residues in these two loops are important for DNA binding but not for protein stability or multimerization (Ueguchi et al. 1996).

Despite having a detailed structure of the H-NS DNA-binding domain and information about which residues interact with DNA, it remains unclear how the interaction occurs between H-NS and DNA including what interactions determine specificity, although footprinting analyses indicate that interactions occur primarily within the major groove (Rimsky and Spassky 1990; Tippner et al. 1994). Mutational analysis on the domain has identified several residues critical for DNA binding. Notably, a mutation in P115 leads to a molecule that has lost its ability to distinguish between a model curved sequence and a non-curved sequence suggesting there may be two modes of binding (Spurio et al. 1997). Multiple binding modes for H-NS have also been proposed as the result of two different biophysical studies, but no details about exactly how the two binding modes specifically differ have been determined (Shindo et al. 1999; Tippner and Wagner 1995). Even without a clear model of the specific binding mechanism a sufficient amount of progress has been made to propose a rational working model for how H-NS binds to specific genes and how it regulates transcription. The model and the data supporting it are outlined in the following section.

13.4 Binding of H-NS to DNA

Much of our current understanding about how H-NS interacts with DNA is derived from studies of several different H-NS regulated promoters that employed standard techniques to measure protein-DNA interactions including electrophoretic mobility shift (EMSA, gel-shift) assays and footprinting analysis. However recent advances in both single molecule analysis (e.g. atomic force microscopy and single molecule manipulation) as well

as high-throughput analysis of H-NS binding sites by chromatin immunoprecipitation and microarray analysis have greatly expanded our working knowledge of which sequences are H-NS targets and how binding of H-NS generates higher-order nucleoprotein complexes.

DNAse I protection assays performed by several laboratories studying different loci have revealed that H-NS binds to multiple sites in cooperative fashion (i.e. multiple sites are occupied simultaneously in a very narrow concentration range) (Donato et al. 1997; Falconi et al. 1998; Rimsky and Spassky 1990). Footprints often extend well beyond what would be expected for a single bound protein dimer, indicating that H-NS may be capable of generating higher-order complexes on DNA. The vast majority of these footprints cover regions that are very AT-rich and a surprisingly large number of them are found quite distant from the -35 and -10 core promoter elements (Nagarajavel et al. 2007; Zuber et al. 1994). Studies of individual loci as well as high-throughput chromatin immunoprecipitation studies have found that the regions bound by H-NS are often well upstream or downstream of the promoter start site and appear quite often within the coding region of the gene (Dattananda et al. 1991; Grainger et al. 2006; Lucchini et al. 2006; Navarre et al. 2006; Olekhnovich and Kadner 2006; Overdier and Csonka 1992; Owen-Hughes et al. 1992; Schnetz 1995; Wolf et al. 2006; Yang et al. 2005).

It is unclear what specific feature of AT-rich sequences is recognized by H-NS. Many groups have proposed that H-NS targets intrinsically curved DNA, specifically DNA with planar curvature (Owen-Hughes et al. 1992; Prosseda et al. 2004; Tupper et al. 1994). Indeed, H-NS was identified in a biochemical screen as a DNA binding protein that displayed a marked preference for a model curved sequence over a non-curved sequence (Yamada et al. 1990). However, a genome-wide analysis of H-NS binding sites by chromatin immunoprecipitation found a much stronger correlation between H-NS binding and AT-content than with curvature (Lucchini et al. 2006) – specifically that AT-rich regions were 20 times more likely to be bound by H-NS than GC-neutral or GC-rich regions whereas curved regions were only twice as likely to be bound as regions predicted to have no curvature. Lucht et al. found that H-NS bound strongly to an extended region at the 5' end of the *proV* gene despite the fact it contained no measurable curvature (Lucht et al. 1994) and in other studies H-NS binding has been difficult to demonstrate at some regions predicted to be curved (Jordi et al. 1997; O'Gara and Dorman 2000). A highly curved synthetic DNA sequence (Ulanovsky et al. 1986) employed in a number of other H-NS binding studies is also incidentally extremely AT-rich (TCTCTAAAAAATATATAAAAA; %GC = 9.5) (Zuber et al. 1994), which further confounds interpretations whether AT-richness or curvature is the critical feature being recognized. It is clear that the AT-rich sequences bound by H-NS are highly variable, and the affinity of H-NS for some *bona fide* sites (i.e. sites bound by H-NS in vivo) differs from that of non-specific sites by less than an order of magnitude (Lucht et al. 1994; Tupper et al. 1994). Addition of high concentrations of H-NS to a binding reaction can lead to a DNA-molecule that is almost completely coated with H-NS (Dame et al. 2000; Schneider et al. 2001). Indeed, overexpression of H-NS in vivo can lead to a complete compaction of the chromosome, resulting in cell death (Spurio et al. 1992).

The fact that H-NS binds such a wide variety of sequences led to the widespread belief that H-NS lacks a “consensus” recognition motif typical for most proteins that interact with DNA. This view has been challenged by the recent discovery of a defined 10 nucleotide high-affinity binding site in the *proV* promoter (Bouffartigues et al. 2007). H-NS displays remarkable affinity for this site with an apparent K_D of approximately 15 nM. This small binding motif can be introduced into other sequences including GC-rich regions and retain much of its affinity indicating that this motif is alone sufficient for H-NS recognition. A bioinformatic analysis determined that sequences highly similar to this motif are statistically overrepresented in many regions bound by H-NS in vivo as determined by genome wide chromatin-immunoprecipitation studies (Lang et al. 2007). These high affinity sites may serve as nucleation sites for H-NS binding, which facilitate the subsequent recruitment of other H-NS molecules to adjacent or nearby sites of lower affinity.

In addition to its ability to act cooperatively to form extended nucleoprotein filaments on target DNA, H-NS and related molecules also display another unusual characteristic that is important for their function. Atomic (or scanning) force microscopy has revealed that H-NS and related proteins are capable of “bridging” DNA strands; suggesting that the DNA binding subunits of each protomer are arranged in a manner that orients them on opposing faces of the H-NS dimer (Dame et al. 2000, 2001, 2002, 2005). These mechanistic distinctions may seem trivial at first but they are actually quite powerful in explaining much of what we know about how H-NS silences transcription, condenses DNA, and constrains supercoils (Fig. 13.2).

Advances in the ability to manipulate individual DNA molecules in complex with proteinacious factors have provided insights into H-NS function that could not be easily addressed using traditional approaches. A recent study performed by Dame, Noom, and Wuite provided important data regarding the H-NS/DNA interaction (Dame et al. 2006). In this study the ends of two individual bacteriophage lambda DNA molecules were each tethered to micrometer sized polystyrene beads. Each of these beads could be manipulated by means of a specialized instrument (the “Q-trap”) with four independently movable optical traps. This setup allowed force to be applied to a bridged DNA/H-NS/DNA complex by pulling on the DNA ends in directions that result in either “unzipping” or “shearing” of the complex (Fig. 13.3). The measurements obtained while unzipping an H-NS bound region indicate that the spacing of each barrier (a single H-NS binding event) is highly variable and sometimes corresponds to a size of one H-NS DNA binding domain. This suggests that the fundamental unit of binding/bridging is likely to be a dimer and that the apparent cooperativity between H-NS molecules as they polymerize along an AT-rich patch is not due to interactions between H-NS molecules per se, but rather that AT-rich DNA duplexes are brought into proximity with one another which facilitates the binding of additional H-NS subunits to lower affinity sites. This observation conflicts somewhat with the earlier studies indicating that the functional unit of H-NS binding is a tetramer or larger multimer (Esposito et al. 2002; Spurio et al. 1997); leaving the question of whether higher-order H-NS complexes play a role in gene regulation unresolved.

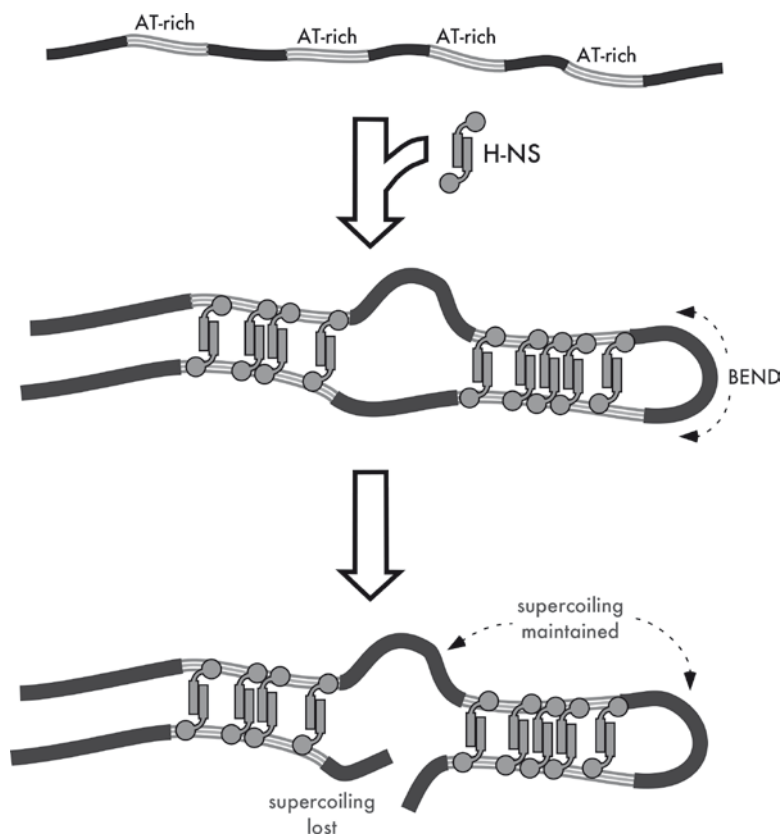


Fig. 13.2 DNA bridging by H-NS explains how curvature facilitates binding and how H-NS may function as a topological barrier. The affinity of H-NS for AT-rich DNA (*striped*) and its ability to polymerize along and bridge adjacent regions of DNA provides an explanation for several previous observations. The H-NS dimer has diametrically opposed DNA binding surfaces. Intrinsic curvature or a protein-induced bend (BEND) can serve to arrange nearby AT-rich sites to facilitate H-NS binding. This model can explain why many H-NS binding sites lie downstream of the promoter and provides an explanation as to why H-NS displays higher affinity for repeating A-tracts on the same face of the double helix *in vitro* (an arrangement that also induces a bend). Furthermore this arrangement can account for the ability of H-NS to constrain supercoiling to local domains: a nick or break in one domain would not alter supercoiling in an adjacent topologically isolated domain (*bottom*). Note that this figure is not drawn to scale and that domains in the chromosome in between H-NS patches are frequently much larger than depicted here

Measurements with the Q-trap also revealed several other H-NS binding parameters that were difficult to determine with more traditional approaches (Dame et al. 2006). First, it was found that H-NS dimers create a barrier to transcription of approximately 7 pN, which is relatively weak compared to the force generated by translocating RNAP (25 pN) (Wang et al. 1998). Second, the Q-trap enabled the measurement of off-rates, which were measured to be relatively fast for each DNA binding domain in a dimer. The combination of cooperativity and a high off-rate

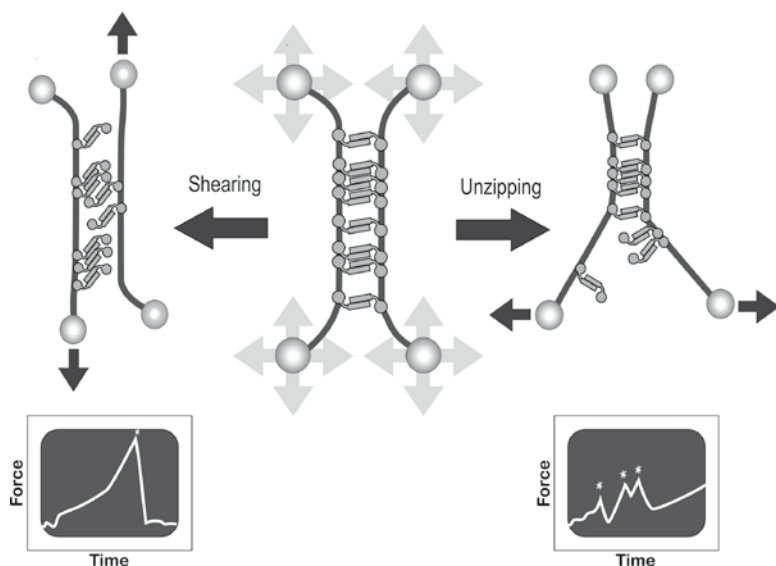


Fig. 13.3 The Q-trap experiment can determine several biophysical parameters for the DNA/H-NS/DNA complex. Polystyrene beads tethered to the ends of two linear DNA molecules bridged by H-NS can each be manipulated by means of optical traps. Application of forces either along the length of the DNA (*left*) or perpendicular to the helix length (*right*) result in either shearing or unzipping of the complex as shown. Readouts of the resulting forces, diagrammed on the monitors below each experiment, can provide information about the nature of the bridging complex as well as the binding forces exerted by each H-NS dimer. In the shearing experiment, movement of the beads leads to increasing resistance until a single catastrophic rupture of the DNA/H-NS/DNA complex occurs. Unzipping leads to a series of small ruptures that can be measured as individual spikes (*three shown below*) in the force landscape

indicated that H-NS bound stretches are stable in their overall structure, yet are able to “breathe” to enable interactions with competing molecules as well as to enable processes like DNA replication and transcription.

Together the data generally support a model whereby H-NS initially nucleates at target sequences for which it has high-affinity (e.g. the *proV* consensus motif). After the initial nucleation event additional H-NS molecules can be recruited to form extended filaments along lower affinity (AT-rich) sites and also bridge adjacent helices. This “bind, bridge, and spread” model not only has strong experimental support but it also has the power to explain several previously perplexing observations regarding H-NS. As outlined in Fig. 13.2 the current model provides a plausible explanation for the observed preference of H-NS for curved DNA (or its higher affinity for regions with AT bases on the same face of the helix) (Bracco et al. 1989; Owen-Hughes et al. 1992; Yamada et al. 1990), its ability to constrain supercoiling into distinct domains (Hardy and Cozzarelli 2005; Hulton et al. 1990), its ability to facilitate DNA bending (Spurio et al. 1997), its ability to affect gene expression by binding regions downstream of the transcription start site (Atlung and Ingmer 1997; Chen et al. 2005; Dole et al. 2004b; Jordi and Higgins 2000;

Owen-Hughes et al. 1992; Yang et al. 2005) and the cooperative interaction of two distantly spaced H-NS binding sites (Donato et al. 1997; Falconi et al. 1998; Nagarajavel et al. 2007; Pul et al. 2008; Rimsky and Spassky 1990).

Under certain conditions DNA bridging is not the only way in which H-NS binds to DNA. In the single molecule manipulation experiments performed by Dame et al., it was possible to pre-coat individual strands with H-NS and prevent their mutual interaction via bridging. One explanation is that the AT-rich sites on each strand were bound by H-NS dimers in such a manner that one binding site of each dimer was left unoccupied. Since both strands would be coated in this fashion there would be no sites available on each strand to allow bridging between them. Another possible explanation is that H-NS dimers may be able to bind with both DNA-binding subunits oriented toward the same strand. This experiment also indicates that bridging does not occur through protein-protein interactions between bound dimers, but rather that adjacent DNA is linked through interactions with each dimer. The possibility that, under certain conditions, H-NS can bind DNA with each of its subunits interacting with the same strand (the “non-bridging mode”), could theoretically explain discrepant results in DNA binding observed in single molecule studies performed by Amit and Stavans where H-NS/DNA interactions were analyzed using magnetic tweezers whereby the addition of H-NS to single DNA molecules was found to extend and stiffen the DNA with no evidence of bridging, which would be expected to compact the molecule (Amit et al. 2003; Amit et al. 2004; Dame and Wuite 2003).

The bridging model of H-NS binding suggests that H-NS binding can be facilitated by curvature without recognizing curvature per se. The correlation with curvature may in part reflect the fact that the DNA arms flanking the curve are brought within proximity to each other, an arrangement that would facilitate bridging by lowering a significant entropy barrier. This model has been invoked to explain the ability of H-NS to trap RNAP in a loop at the *rrnB* promoter (Dame et al. 2002), the *hdeAB* promoter (Shin et al. 2005) and the *virF* promoter (Prosseda et al. 2004). This model can also explain why distamycin, a drug that binds the minor groove and alters curvature, can disrupt H-NS binding at a model curved AT-rich sequence (Yamada et al. 1990).

13.5 H-NS as a Negative Regulator of Foreign-Derived (Xenogeneic) DNA

Across the bacterial kingdom surprising plurality of sequences obtained via LGT is AT-rich when compared to their resident genomes (Daubin et al. 2003b). For the enteric bacteria this includes the vast majority of virulence-associated sequences (i.e. pathogenicity islands and islets) as well as several genes involved in antibiotic resistance, almost all of which are xenogeneic (derived from a foreign source). The ability of H-NS to selectively bind and silence AT-rich sequences makes it the major virulence regulator in *E. coli* and *Salmonella* as well as in other bacterial species including the *Yersiniae* and the *Vibrios*. Accordingly, H-NS is essential for

virulence in *Salmonella* (Harrison et al. 1994) and uropathogenic *E. coli* (Muller et al. 2006) and it also plays a role in drug resistance by negatively regulating the expression of plasmid encoded β -lactamase and multiple drug pumps in *E. coli* as well as the cryptic *eefABC* multidrug resistance pump in *Enterobacter* sp. (Masi et al. 2005; Nishino and Yamaguchi 2004; Zuber et al. 1994).

Two separate studies have employed genome-wide cDNA microarray analysis to identify genes that display altered expression in *hns* mutants of *S. Typhimurium* (Lucchini et al. 2006; Navarre et al. 2006). In accordance with previous estimates, increased expression of more than 400 genes was documented in an *hns* mutant. Most of these H-NS silenced genes were AT-rich; the average GC-content of an H-NS silenced ORF was 46.8%, in comparison to 52.2% for the overall genome. Approximately 90% of the H-NS silenced genes showed evidence of being acquired by LGT, and 65% of H-NS silenced genes were unique to *Salmonella* spp., indicating the major role that H-NS plays in regulating expression from xenogeneic genes. Furthermore H-NS could bind and silence the expression of an AT-rich gene from *Helicobacter pylori* that was engineered along with its promoter into a large non-essential GC-neutral region of the *Salmonella* genome. As predicted, H-NS did not target the adjacent GC-neutral regions as determined by chromatin immunoprecipitation. This finding demonstrated that H-NS silences newly introduced sequences based on increased adenine and thymine content per se and irrespective of their position on the chromosome.

Both of the studies mentioned above also performed H-NS chromatin immunoprecipitation combined with microarray analysis ("ChIP-on-chip") to determine the binding sites of H-NS throughout the *Salmonella* chromosome (Lucchini et al. 2006; Navarre et al. 2006). This approach revealed a striking correlation between %AT content and H-NS binding with the magnitude of interaction with H-NS corresponding closely with the degree of local AT-content. More than 400 AT-rich regions of the *Salmonella* chromosome, including the plasmid virulence region, all five pathogenicity islands and nearly every AT-rich islet, were bound by H-NS. A high-resolution oligonucleotide array used in one study (Navarre et al. 2006) demonstrated that binding is not necessarily restricted to promoter regions, as several AT-rich coding regions were also bound by H-NS. These findings are consistent with what has been previously noted in the *bgl*, *proU* and *eltAB* operons (Dole et al. 2004a; Lucht et al. 1994; Madhusudan et al. 2005; Yang et al. 2005). By using identical growth conditions for both expression analysis and chromatin immunoprecipitation, Lucchini et al. concluded that H-NS acts almost exclusively as a silencer of gene expression and that almost all cases of gene activation by H-NS were due to indirect mechanisms (Lucchini et al. 2006).

Two parallel studies that employed ChIP-on-chip technology to determine the binding sites for H-NS in *E. coli* came to conclusions similar to those of the studies in *Salmonella* (Grainger et al. 2006; Oshima et al. 2006). Namely, H-NS primarily targets AT-rich sequences and, although biased toward intergenic regions, frequently binds within coding sequences. Like Lucchini et al., Oshima et al. found that H-NS acts primarily as silencer of transcription as most H-NS bound loci were transcriptionally downregulated. Earlier attempts to characterize the H-NS regulon in *E. coli* K-12 had been performed using proteomic or cDNA "macroarray" analysis

(Hinton et al. 1992; Hommais et al. 2001). In the study by Hommais et al. approximately 5% of the *E. coli* transcriptome was altered in the *hns* mutant and many of the genes identified were involved in stress adaptation or cell envelope biogenesis. Although a correlation with foreign DNA was not noted, several genes acquired by LGT were found to be H-NS regulated, including several adhesion loci. A correlation between H-NS binding and foreign genes may have been overlooked because the K-12 laboratory strain lacks obvious large LGT-derived genomic regions such as pathogenicity islands, and the experiments were performed before many related genome sequences were available. More recently, genes under the control of H-NS were determined in a uropathogenic strain of *E. coli* (UPEC) using an expanded “pathoarray” designed specifically to analyze genes involved in virulence (Muller et al. 2006). In this study every UPEC virulence locus was found to be downregulated by H-NS, strongly corroborating the findings in *Salmonella* and indicating that H-NS function as a master virulence regulator is conserved between the two species.

Mutations in *hns* are lethal in *Yersinia* and this fact has slowed the characterization of H-NS and its role in these bacteria (Ellison and Miller 2006; Heroven et al. 2004). Banos et al. have recently exploited the dominant negative properties of truncated H-NS like molecules that contain only the dimerization domain to characterize the genes under control of H-NS in *Y. enterocolitica* using 2D-gel electrophoresis (Banos et al. 2008). Using this technique they found that H-NS plays a role in regulating the expression of *proV*, *ureG*, *galU* and the *ymoA* genes similar to what has been observed in other enteric bacteria. Their study most certainly underestimated the genes under control of H-NS in *Yersinia* as they did not identify several genes previously known to be silenced by H-NS. However their approach may prove useful to determine the complete set of genes under control of H-NS when coupled to microarray analysis.

The phenomenon of downregulation of genes acquired by foreign sources on the basis of their lower GC-content has been termed xenogeneic silencing (i.e. the silencing of sequences derived from foreign sources) (Navarre et al. 2006, 2007). Xenogeneic silencing might explain many perplexing observations regarding bacterial genomes in the context of LGT. It seems likely that H-NS has facilitated the acquisition of AT-rich genes by allowing them to be tolerated better than they would otherwise be if their expression levels were unmitigated. Indeed, *hns* mutations are associated with fitness defects that range from somewhat mild in *E. coli* to severe in *Yersinia*, presumably because of the simultaneous misregulation of dozens to hundreds of xenogeneic genes. The fact that the *Yersinia* do not possess a second H-NS paralogue that can partially compensate for a loss of H-NS may account for its essentiality. A similar fitness defect has been observed in *Pseudomonas* where deletion of *mvaT* or *mvaU*, genes encoding two H-NS like proteins, does not result in a strong phenotype while strains with mutations in both genes cannot be constructed (Castang et al. 2008). The degree to which H-NS-like molecules enhance fitness is species specific for reasons that remain unclear; both H-NS paralogues in *E. coli*, StpA and H-NS, can be deleted with only a moderate decrease in fitness compared to the single *hns* mutant (Sondén and Uhlin 1996).

Additional support for the model that H-NS plays a primary role in mitigating the untoward effects that can result from high-level expression of xenogeneic

sequences can be found by examining the compensatory mutations that increase growth rate of *hns* mutants. Deletion of *rpoS* and *phoP*, two genes that positively regulate the expression of many xenogeneic loci, or deletion of SPI-2 in each can lead to a significant enhancement of growth rate in *Salmonella hns* mutants (Lucchini et al. 2006; Navarre et al. 2006). A similar enhancement in fitness was previously observed when *rpoS* mutations were introduced into *hns* mutants in *E. coli* (Barth et al. 1995). It has been observed that *E. coli hns* mutants have unstable genomes that can spontaneously undergo large deletions, presumably due to the increase in fitness that results from the loss of certain detrimental H-NS repressed loci (Lejeune and Danchin 1990). Interestingly, during the course of studies on *Salmonella hns* mutants by our group we noticed that one of our mutant strains spontaneously lost a region of the genome near the strongly H-NS dependent *pagC* gene (see supplementary data of Navarre et al. (2006)). Together these data support a hypothesis that it is misregulation of specific gene sets, rather than a gross alteration of chromosome structure per se, that causes the majority of the observed fitness defect in *hns* mutants. This idea, however, remains to be formally tested.

13.6 The Hha/YmoA Family of Accessory Regulators

At many loci silencing by H-NS is augmented by the Hha-like proteins, a family of small (~8 kDa) soluble proteins that, like H-NS, are critical for virulence and cause a number of seemingly disparate phenotypes when mutated (Coombes et al. 2005; Madrid et al. 2007a). The prototypical member of this family, Hha, was originally identified as a negative regulator of the plasmid encoded alpha-haemolysin gene, *hlyA*, of some pathogenic strains of *E. coli* (Godessart et al. 1988; Nieto et al. 1991). The *Yersinia* orthologue of Hha, called YmoA, was simultaneously identified as having a negative regulatory effect on the important virulence regulator VirF in *Yersinia* and later as a regulator of several other *Yersinia* virulence factors including invasins (Cornelis et al. 1991; de la Cruz et al. 1992; Ellison et al. 2003). Most isolates of *E. coli*, *Shigella* and *Salmonella* possess at least one additional Hha paralogue, YdgT (sometimes called Cnu), and some strains carry additional paralogues encoded on mobile genetic elements like conjugative plasmids (Forns et al. 2005).

Members of the Hha-like proteins share significant functional redundancy (Nieto et al. 2002). Hha paralogues, like YmoA, YdgT and the ORF182 from the R27 plasmids can complement *E. coli hha* mutants for downregulation of several loci and, reciprocally, Hha can functionally replace YmoA in *Yersinia* mutants although the degree of complementation observed depends on copy number (Balsalobre et al. 1996; Forns et al. 2005; Mikulskis and Cornelis 1994; Nieto et al. 2002; Paytubi et al. 2004). Because of their functional redundancy single mutations in either *hha* or *ydgT* in *E. coli* or *Salmonella* are well tolerated and do not result in an obvious growth defect. At loci that display regulatory perturbation in *hha* or *ydgT* single mutants, the effect is almost always more severe for the *hha* mutation than for the *ydgT* mutation (Paytubi et al. 2004; Silphaduang et al. 2007).

In all cases tested thus far, regulatory perturbation at such loci is exacerbated in the double *hha ydgT* mutant to a degree that is considerably greater than the additive contributions of both single mutants (Paytubi et al. 2004; Silphaduang et al. 2007). Together these results indicate that the Hha-like molecules are highly redundant with regard to their function and that, in wild-type *E. coli* and *Salmonella*, Hha is the dominant functional effector. YdgT apparently plays a secondary, or “backup”, role at several loci that only becomes apparent in the absence of Hha.

Recently microarray analysis has been used to determine the regulon under control of YdgT and Hha in *S. Typhimurium* (Vivero et al. 2008). Transcriptome analysis of *hha* and *ydgT* single mutants were not reported but the *ydgT* and *hha* double mutant strain displayed dramatic and genome-wide misregulation of over 1,000 genes with a pattern indicating that Hha and YdgT together play an important role in silencing xenogeneic AT-rich sequences. Further supporting this model is an early study finding that mutations in *hha* led to enhanced expression of a recombinant highly AT-rich sequence encoding an endoglucanase from *Clostridium cellulolyticum* (Blanco et al. 1991). The majority of the genes downregulated by Hha and YdgT in the microarray study show evidence of having been acquired via LGT and share significant, but incomplete, overlap with the genes under control of H-NS (Navarre et al. 2006; Vivero et al. 2008). These include genes from SPI-1, 2, 3 and 5 as well as a large number of other small genomic islets. A very large number of genes also displayed lower expression in the *hha ydgT* double mutant; among those are genes involved in motility, secretion, and the generation of surface structures. The microarray study also revealed that some loci directly silenced by H-NS, including *proU*, are unaffected by mutations in *hha* and *ydgT* (Vivero et al. 2008). Although only one growth condition was examined in this study the findings bring up the question of whether all or only a subset of H-NS-regulated loci require Hha or YdgT as cofactors for silencing. Current data suggest that the levels of Hha/YdgT are much lower than the levels of H-NS under most growth conditions suggesting that they cannot act in a 1:1 (or 1:2, etc) complex at all promoters; rather it suggests that Hha and YdgT may selectively bind H-NS only at certain promoters and not others. Clarification of the role that Hha and YdgT play at H-NS regulated promoters will be greatly facilitated by determining the genome-wide subset of loci that associate with Hha or YdgT using ChIP-on-chip assays.

The mechanism by which Hha-like molecules affect gene expression at H-NS regulated loci remains unclear and there is considerable debate as to whether the Hha-like proteins exert their actions by binding DNA directly or through an interaction with H-NS. It is possible that Hha and YdgT interact both with DNA and with H-NS, helping H-NS to specifically regulate a subset of promoters in a manner analogous to the way in which the extradenticle protein of *Drosophila* directs the specificity of Hox proteins to certain promoters (Joshi et al. 2007). Experimental support for the model that Hha exerts its regulatory effects as a DNA binding protein is derived from several observations. Hha and YdgT have been found to bind DNA fragments in gel-shift and footprint assays and also in less direct genetic assays for DNA binding (Kim et al. 2005; Olekhnovich and Kadner 2007).

Furthermore, mutations in *hha* can affect gene expression in the absence of H-NS, suggesting that Hha is independently capable of interacting with promoters. The relevance of these findings has been called into question as other studies have found that Hha binds DNA with relatively low affinity and with poor ability to target specific sequences over non-specific competitors including poly(dI-dC) or salmon sperm DNA (Ellison and Miller 2006; Madrid et al. 2002a; Nieto et al. 2000, 2002). Given that Hha and H-NS can co-purify (see below) there is also some concern that Hha preparations used in some in vitro binding studies were not sufficiently free of H-NS. Furthermore, the studies in *E. coli* finding that Hha can affect gene expression in the absence of H-NS were performed in cells that maintained wild-type expression of the H-NS paralogue StpA, through which Hha could alternatively exert its effects.

A growing body of evidence supports the idea that Hha-like molecules affect gene expression via H-NS, in particular through specific interactions with the H-NS N-terminal dimerization domain. Hha, YdgT and YmoA have each been shown to interact with H-NS in vitro either through pull-down assays or by co-purification (Nieto et al. 2002). Indeed one method to prevent Hha from entering inclusion bodies when overexpressed from recombinant expression vectors in *E. coli* is through co-expression with recombinant H-NS (Pons et al. 2004). YmoA is also capable of making a complex with H-NS at the *inv* promoter that is distinct from the complex formed by either molecule alone (Ellison and Miller 2006). The solution structures of Hha, YdgT and YmoA have been determined and each are composed entirely of alpha-helices with similar, but not identical, arrangements. YdgT is composed of three helices while Hha and YdgT have an additional small C-terminal helix (Bae et al. 2008; Garcia et al. 2005; McFeeters et al. 2007; Yee et al. 2002). Spectral peaks corresponding to residues in Hha are altered in ^1H - ^{15}N HSQC-NMR spectra upon addition of H-NS and, in reciprocal experiments, it was found that HSQC spectra of H-NS are altered by the addition of Hha (Garcia et al. 2005, 2006). The structural changes that occur in Hha upon addition of H-NS result from a change in the tertiary fold of the molecule (as opposed to changes in the secondary structure) and mimic, in part, changes in Hha that are caused by an increase in temperature (Garcia et al. 2005). The residues affected by H-NS binding are displayed in Fig. 13.4.

Random mutagenesis has revealed that Hha residues R16, R50, and P64 are critical both for Hha function and for its ability to interact with H-NS (Nieto et al. 2002). The ^1H - ^{15}N HSQC spectra of these three residues were not significantly perturbed by the addition of H-NS although they lie on the same face of the Hha protein as the residues that did display significant alterations in their spectra (Fig. 13.4). Recently another mutation has been identified in Hha (C18I) that abrogates the ability of Hha to downregulate gene expression at the *hlyA* locus in vivo without significantly affecting its overall structure or its ability to bind H-NS in vitro, in fact binding to H-NS appears to be slightly stronger with the C18I mutant (Cordeiro et al. 2008). This finding suggests that the ability of Hha to bind H-NS per se is insufficient for function. Instead it supports a model whereby Hha has an effect on the higher order structure of nucleoprotein complexes that is perturbed

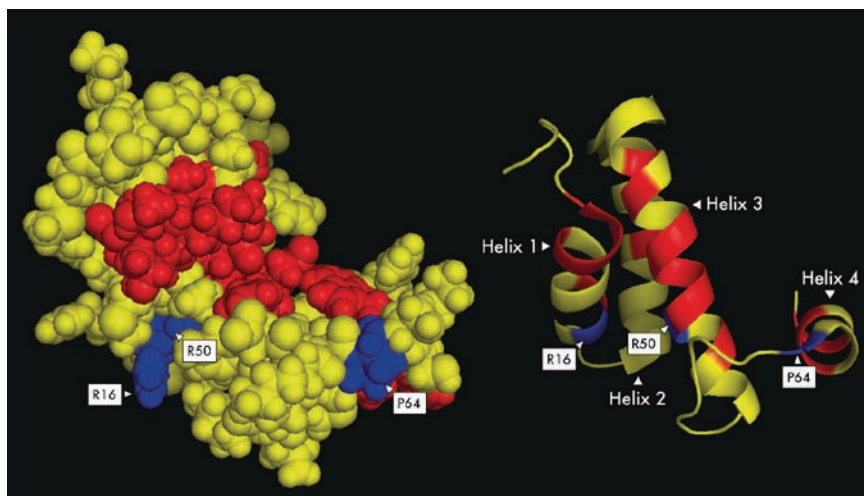


Fig. 13.4 The solution structure of Hha. The solution structure and results of the 1H - ^{15}N HSQC experiment of H-NS interactions with Hha are shown (Garcia et al. 2005; Yee et al. 2002). The models are displayed either as a space filling (*left*) or ribbon diagram (*right*) to facilitate the visualization of residue proximity and tertiary structure. Residues in Hha that are perturbed upon binding to H-NS are highlighted in lighter grey. Residues R15, R50 and P64, which are essential for the interaction of Hha with H-NS are labeled and shown in white

by the C18I mutation. Interestingly, the directed replacement of this cysteine with either alanine or serine did not recapitulate the phenotype of the isoleucine replacement indicating that the chemistry of the cysteine sulfhydryl group is dispensable for Hha function.

The ^1H - ^{15}N HSQC NMR study to determine the H-NS regions responsible for interacting with Hha identified key residues as residing within helices 1 and 2 as well as the intervening loop (Garcia et al. 2006). In particular, ^1H - ^{15}N HSQC spectral peaks corresponding to H-NS residues 7 through 14 were severely broadened upon the addition of Hha. Mutagenesis of residue R12 was shown to dramatically reduce the interaction of H-NS with Hha. A nearby highly conserved arginine, R15, was found to play a role in H-NS dimerization but did not display a strong effect with regard to the ability of H-NS to bind Hha. Both NMR-based studies of the Hha/H-NS interaction support a relative stoichiometry of one Hha molecule per two H-NS monomers (or one H-NS dimer) (Garcia et al. 2005, 2006).

The Hha-like molecules share small but significant sequence identity with the N-terminal domain of H-NS with a highly conserved stretch of residues between the beginning of helix 3 and the preceding loop of H-NS and the beginning of helix 2 and the loop preceding it in Hha/YmoA (Garcia et al. 2006; McFeeters et al. 2007; Nieto et al. 2002). It has also been shown that helices 1 and 2 of YmoA and Hha share a similar structural arrangement to helices 2 and 3 in H-NS (Garcia et al. 2006; McFeeters et al. 2007). Furthermore it has been shown that replacement of the N-terminal 64 residues of H-NS with residues 1 through 60 of Hha generates a

molecule that can partially compensate for function in an *E. coli hns* mutant (Rodriguez et al. 2005). These observations led to the hypothesis that Hha-like molecules may function by intercalating into higher-order H-NS complexes on DNA by substituting for an H-NS dimer (McFeeters et al. 2007). This model is also supported by the observation that the N-terminal domain of H-NS can form higher-order oligomers at lower concentrations in the presence of Hha than are needed to form oligomers in the absence of H-NS (Garcia et al. 2005). Although the hypothesis that Hha-like molecules are analogous to H-NS dimerization domains is attractive in many respects it may be somewhat over-simplistic. First, H-NS dimerization domains, when overexpressed, act in a dominant negative fashion to antagonize silencing rather than augment it (Ueguchi et al. 1997). Second, residues of H-NS that are involved in its interaction with Hha are proximal to, but not identical to, the residues within H-NS involved in dimerization and Hha binding does not alter the ^1H - ^{15}N HSQC spectra of residues involved in dimerization (Garcia et al. 2006). Therefore, if the Hha-like molecules do intercalate into higher-order H-NS complexes at promoters, it is clear that they do so in a manner that is distinct from the way in which H-NS monomers interact with one-another to form dimers.

The Hha-like molecules are found only in the *Enterobacteriaceae* and their conjugative plasmids despite the fact that H-NS-like molecules are found in a large number of Gram-negative proteobacteria outside of the enteric lineage (Madrid et al. 2007b). Species like *E. coli/Shigella* and *Salmonella* that encode two H-NS like molecules (H-NS and StpA) generally also encode the two paralogs, Hha and YdgT, while the *Yersiniae*, which encode only one H-NS only encode one Hha-like molecule (YmoA). Whether this suggests that each paralogue makes specific partnerships, e.g. Hha is the preferred partner for H-NS and YdgT is the preferred partner for StpA, is difficult to determine due to the strong degree of functional redundancy between each of these molecules. Interestingly, the species distribution of Hha parallels the distribution of a specific sequence motif (NNIRTL) that is highly conserved in the H-NS molecules of the *Enterobacteriaceae* but is absent in all other H-NS homologues sequenced to date. The central arginine residue in the motif corresponds to the R12 of H-NS that is critical for Hha binding, providing further support that this motif is highly likely to be relevant for function (Madrid et al. 2007b).

Despite what we now understand about the Hha-family of molecules it remains unclear what the exact roles of the Hha-like molecules are in the enteric bacteria and why bacterial species that encode H-NS-like molecules, like *Vibrio* or *Pseudomonas*, do not encode Hha paralogs. It is possible that Hha-like molecules provide regulatory inputs to control the activity of H-NS-mediated silencing in response to environmental conditions including osmolarity, growth phase or temperature. Indeed, most genes that are Hha/H-NS regulated respond to a change in one of these environmental conditions (Duong et al. 2007; Ellison et al. 2003; Mikulskis et al. 1994; Mourino et al. 1996, 1998; Ono et al. 2005). Also worth mentioning is the role Hha plays in regulating conjugation and transposition. Mutations in *hha* lead to an increase in conjugation while Hha overexpression leads to a decrease in conjugation frequency, an effect that has been observed in more than one family of plasmids (Forns et al. 2005; Mikulskis and Cornelis 1994;

Nieto et al. 1998). Overexpression of Hha leads to an increase in transposition by an undefined mechanism, similar to what has been noted for H-NS (Balsalobre et al. 1996; Mikulskis and Cornelis 1994; Shiga et al. 2001; Swingle et al. 2004). A discussion on the role of H-NS/Hha as an environmental regulator and as a regulator of conjugation and transposition is expanded further below.

13.7 Mechanisms of H-NS-Mediated Transcriptional Downregulation

It was recognized early on that H-NS regulates the expression of a large number of genes in both *E. coli* and *Salmonella*, and that its effects on gene expression are largely, if not exclusively, inhibitory (Dorman 2004; Higgins et al. 1990). Mechanistic analysis of H-NS-mediated silencing has been analyzed in detail at a number of loci including *proU*, *bgl*, *dps*, *hdeAB*, the *rrn* operon, the *virF* locus of *Shigella*, and at the *hns* gene itself. What has emerged from these studies is that H-NS can downregulate gene expression in a number of mechanistically distinct ways that will be outlined in the following sections. Simple categorization of silencing mechanisms is difficult given that several factors including temperature, osmolarity, promoter activity, and other DNA binding proteins coordinate in varied ways to affect expression at each promoter. For all proposed silencing mechanisms, however, the ability of H-NS to bind cooperatively and to generate extended nucleoprotein filaments appears to be important (Dame 2005; Rimsky 2004; Spurio et al. 1997; Ueguchi et al. 1997).

13.7.1 Occlusion

The most obvious way in which H-NS binding could downregulate gene expression is by simply preventing access of RNAP to core promoter elements through steric competition. H-NS binding sites have been shown to overlap core promoter elements at a number of promoters by footprinting analysis, but in such cases it is not necessarily safe to infer that RNAP cannot bind in the presence of H-NS. To date very few promoters have been examined where occlusion of RNAP by H-NS has been formally recapitulated in vitro. H-NS appears to greatly reduce RNAP $\cdot \sigma^D$ (but not RNAP $\cdot \sigma^S$) binding at the *dps* promoter (Grainger et al. 2008) in vitro. H-NS also partially competes for binding of RNAP $\cdot \sigma^D$ to the *E. coli hlyE* promoter (Lithgow et al. 2007). At both the *proU* and *bgl* promoters H-NS appears to block transcription by RNAP at a step before open complex formation but binding/competition assays between RNAP and H-NS have not been performed in vitro (Jordi and Higgins 2000; Nagarajavel et al. 2007) and it is possible that RNAP binds to these promoters but cannot initiate open complex formation.

At promoters where occlusion is occurring, one prediction is that H-NS and RNAP would not be observed to co-associate during chromatin immunoprecipitation assays. Three recent high-throughput chromatin immunoprecipitation studies in *E. coli* and *Salmonella* have mapped the genome wide co-occupancy of H-NS and RNAP at promoters with conflicting results. In the *Salmonella* study it was found that the vast majority of H-NS bound promoters did not co-precipitate with RNAP (Lucchini et al. 2006), while approximately half the H-NS-bound promoters in *E. coli* were shown to co-precipitate with RNAP (Grainger et al. 2006; Oshima et al. 2006). This discrepancy remains to be explained but it is unlikely to be caused by the difference in the species tested given that *E. coli* and *Salmonella* are closely related. Chromatin immunoprecipitation is an assay carried out on a large population of cells and given the heterogeneity of expression in individual cells in bacterial populations it is impossible to determine whether H-NS and RNAP appear to co-localize at a given region in these assays due to their simultaneous occupation in most cells or instead merely appear to co-localize due to ensemble averaging of two distinct subpopulations that might exist within the culture, one population with RNAP bound at a given locus and another where H-NS is bound.

13.7.2 Polymerase Trapping

H-NS can also downregulate gene expression at steps after the association of RNAP with the core promoter by preventing open-complex formation or by preventing the subsequent establishment of a productive elongation complex. This is supported both by genome-wide chromatin immunoprecipitation studies finding that RNAP and H-NS colocalize at several promoters, as well as in vitro studies at individual promoters including *rrnB* and *hdeAB* (Dame et al. 2002; Schroder and Wagner 2000; Shin et al. 2005). This phenomenon, called “trapping” was first demonstrated at the *rrnB* P1 promoter where H-NS cooperatively binds three distinct patches within the promoter region to enhance rather than reduce RNAP binding (Schroder and Wagner 2000). H-NS mediated silencing at *rrnB* P1 is achieved by preventing RNAP escape from the initial open complex into an elongation complex. Importantly, RNAP and H-NS appear to bind cooperatively with each enhancing the binding of the other. This suggests that RNAP may actually be a co-factor that enables H-NS to make a stable complex at the *rrnB* P1 promoter, perhaps by facilitating bridging as first proposed by Dame et al. (Dame et al. 2002).

The possibility that RNAP plays an active role in making a repressive loop with H-NS is further supported by recent mechanistic studies on the *hdeAB* promoter. Shin et al. employed both permanganate footprinting and atomic force microscopy to analyze the mechanism by which H-NS selectively downregulates *hdeAB* when RNAP is associated with the housekeeping sigma factor σ^D (RpoD, σ^{70}) but permits transcription when the RNAP core is complexed with the alternative sigma factor σ^S (RpoS, σ^{38}) (Shin et al. 2005). Permanganate footprinting shows that H-NS does not prevent open complex formation by either $\text{RNAP} \cdot \sigma^S$ or $\text{RNAP} \cdot \sigma^D$.

Images of the *hdeAB* promoter bound to RNAP $\cdot \sigma^S$ show that the DNA exits at a wide angle, perhaps preventing bridging of the adjacent sequences whereas the promoter bound to RNA RNAP $\cdot \sigma^P$ is bound more tightly and exits from the polymerase at a very narrow angle, allowing H-NS to effectively bridge both flanking sequences and forming a repressive loop.

13.7.3 Supercoil Trapping

As alluded to above, H-NS has been shown to constrain supercoils both in vitro and in vivo (Hardy and Cozzarelli 2005; Higgins et al. 1988; Hinton et al. 1992; Mojica and Higgins 1997; Tupper et al. 1994; Zhang et al. 1996), and H-NS has been termed a “domainin”, i.e. a protein that can prevent global unwinding of the chromosome following double stranded breaks by acting as a barrier that constrains supercoiling to local domains (Hardy and Cozzarelli 2005). In an elegant study, *hns* was identified during a genetic screen for mutants capable of modulating reporter genes under the control of the *gyrB* and *lac* promoters, known to be sensitive to the degree of supercoiling (Hardy and Cozzarelli 2005). Although the control of these promoters by H-NS could be direct (through occlusion or RNAP trapping) it is more likely that the effect observed was due to alterations in local supercoiling since *gyrB* is not strongly regulated by H-NS in its native context.

The ability of H-NS to constrain supercoils could enable it to regulate certain promoters that are sensitive to the degree of supercoiling by locking them in either activated or silenced states (Higgins et al. 1990; Mojica and Higgins 1997). Indeed, changes in supercoiling can alter the expression of a few promoters directly regulated by H-NS including the *bgl* and *proU* operons (Higgins et al. 1988; Mukerji and Mahadevan 1997; Schnetz and Wang 1996). However, the models that suggest that H-NS mediated activation and/or downregulation occurs via alterations in supercoiling have been generated with data from *hns* mutants. This fact means we have little clue as to how supercoiling would be altered at specific promoters in the presence of normal cellular levels of H-NS. One model is that a change in supercoiling in response to some environmental condition at certain promoters leads to a reduction in H-NS binding, thereby enabling transcription (i.e. changes in supercoiling occur at a step before silencing is affected). The other model is that alteration of the H-NS-nucleoprotein complex by a specific transcription factor relieves local constraints on supercoiling thereby allowing transcription elongation or initiation by, for example, aiding in melting the promoter region during open complex formation or in promoter clearance (Dorman 2006; Lim et al. 2003).

These models have proven difficult to test experimentally. Activity at the *bgl* promoter is considerably lower on linear templates than supercoiled templates when assayed by in vitro transcription using RNAP with naked DNA templates, but the same promoter is silenced in vivo regardless of the supercoiling state (Schnetz and Wang 1996). At the *proU* locus, the supercoil trapping model does not fully explain how H-NS-mediated silencing is relieved under high osmolarity conditions

and a significant proportion of the osmotic response at *proU* remains even in the absence of H-NS, indicating that the contributions made by H-NS and supercoiling can be separated at the *proU* promoter (Fletcher and Csonka 1995).

The functional significance of H-NS and its ability to trap supercoils may not be limited to transcriptional regulation. H-NS binding appears to provide topological isolation for both the purpose of localizing genomic perturbations in structure and preventing widespread loss of supercoiling after DNA damage. Studies have suggested that the *E. coli* chromosome is organized into topological domains that vary widely in size but average ≈ 10 kb (Deng et al. 2005; Postow et al. 2004). A recent analysis has shown that the distribution of H-NS binding sites, as determined by high-resolution chromatin immunoprecipitation analysis, correlates well with these empirical estimates for topological domain size and distribution (Noom et al. 2007). There is also a correlation between H-NS levels and domain sizes, and together these suggest that H-NS may be a major factor in determining topological domain size in vivo. Further studies will be needed to directly address the effects of H-NS on the topology of specific domains along the chromosome. A more detailed treatise of the effects of H-NS on chromatin structure is provided in Chapter 8 of this book entitled “Nucleoid-Associated Proteins: Structural Properties”.

13.7.4 Is H-NS-Mediated Transcriptional Downregulation Really “Silencing”?

The term “xenogeneic silencing” was chosen to extend a previously defined conceptual framework regarding the unique mechanism by which H-NS downregulates transcription and to further distinguish it from the mechanism of repression employed by “classical” site-specific DNA binding proteins (Goransson et al. 1990; Rine 1999; Yarmolinsky 2000). The use of the term “silencing” could mistakenly be taken to suggest that similarity exists between silencing by chromatin in eukaryotes and the mechanism that underlies H-NS-mediated downregulation of recently acquired genes. As generally defined eukaryotic and bacterial silencing refer to the downregulation of gene expression by non-specific DNA binding proteins that can act at a distance from the core promoter (Goransson et al. 1990; Madrid et al. 2002b). Drawing parallels between xenogeneic silencing and chromatin-mediated silencing in eukaryotes is superficially attractive given that the latter phenomenon is mediated by histones and that H-NS has been referred to as the bacterial equivalent to a histone.

It is important, however, to reiterate that H-NS-mediated downregulation shares no mechanistic similarity to eukaryotic silencing of gene expression. Unlike silencing mediated by histones, for example, silencing by H-NS does not involve post-translational modification of H-NS by phosphorylation, acetylation or methylation. The effect of H-NS also varies greatly from promoter to promoter; some H-NS-regulated

genes display significant basal levels of expression and therefore are not “silent”, as most eukaryotic genes are. Therefore, as is the case for the term “repression” (the mechanism of which also differs significantly between eukaryotes and bacteria), the silencing terminology only implies a loose functional analogy between bacteria and eukaryotes.

13.8 H-NS as a Temperature and Osmolarity Sensor?

H-NS has been described as a global regulator that can alter the expression of a large number of genes in response to certain environmental conditions such as pH, temperature, anaerobiasis, or osmolarity (Amit et al. 2003; Atlung and Ingmer 1997; Atlung et al. 1996; Dorman 2007). Expression of H-NS relative to DNA content appears to be relatively constant under a range of environmental conditions (Atlung and Ingmer 1997; Dorman 2004; Free and Dorman 1995). Since H-NS concentrations do not vary significantly in response to pH, temperature or osmolarity, the H-NS protein has instead been postulated to undergo structural and functional alteration under these environmental conditions, leading to a corresponding increase in expression at certain H-NS silenced loci. While some biochemical evidence for such structural changes has been obtained, the correlation with H-NS dependent gene expression has been less than perfect. For example, the subsets of genes regulated by pH, temperature, oxygen and osmolarity are distinct; e.g., H-NS silenced genes expressed during conditions of increased osmolarity do not generally overlap with those regulated by temperature (Atlung and Ingmer 1997; Hommais et al. 2001).

Much effort has focused on the possibility that H-NS might act as a thermosensor that globally regulates a large subset of genes in response to temperature. There is some evidence to support this notion including the fact that the DNaseI footprint of H-NS at the *proU* promoter is altered in response to temperature (Badaut et al. 2002). A recent report noted that more than three-quarters of the 531 *Salmonella* genes exhibiting altered expression when cultures were shifted from 25°C to 37°C are dependent on H-NS (Ono et al. 2005). Among these genes were the *Salmonella*-specific SPI-1 and cobalamin biosynthetic (*cob*) genes, as well as genes involved in flagellar biosynthesis and chemotaxis. Both SPI-1 and *cob* were constitutively expressed at high levels in the absence of H-NS regardless of temperature, whereas motility genes were constitutively repressed. Temperature-induced regulation of SPI-1 was rapid, occurring within minutes. Further work found that the purified H-NS N-terminal domain had altered oligomerization properties in response to temperature, and while full-length H-NS had decreased affinity for the *hilC* promoter at 37°C no temperature dependence in the affinity of the C-terminal domain for DNA was observed. Based on these data a model was proposed in which the oligomerization properties of H-NS change at 37°C to favor a dimeric conformation, perhaps through temperature-induced changes in the orientation of the dimerization domain.

Such changes would reduce or alter the mode of DNA binding and permit a rapid transcriptional response to change in temperature.

However, although most temperature-dependent *Salmonella* genes were found to be regulated by H-NS, the converse was not the case. More than 200 genes were silenced equally well by H-NS at 25°C and 37°C, including *virK*, *pipB2*, *mig-14*, *pagC*, *yciEFG*, and the pathogenicity islands SPI-2, SPI-3, and SPI-5. The H-NS-silenced *proU* operon was paradoxically more effectively silenced at 37°C than at 25°C. Furthermore the H-NS-silenced *invA* and *rovA* genes of *Yersinia* are active at 30°C and inactive at 37°C under laboratory conditions (Ellison and Miller 2006; Heroven et al. 2004). Several loci in the locus of enterocyte effacement (LEE) carried by pathogenic *E. coli* are silenced by H-NS at all temperatures (Umanski et al. 2002). H-NS silences *eltAB* encoding the heat-labile enterotoxin of *E. coli* both at 22°C and at 37°C (Yang et al. 2005). Thermal regulation of the *Shigella virF* promoter has been attributed to temperature-induced alterations in the conformation of a bend adjacent to the promoter rather than in the H-NS protein itself (Falconi et al. 1998; Prosseda et al. 2004). Changes in promoter conformation might also account for the persistent silencing of the *virB* promoter at 37°C when the *Shigella* virulence plasmid is integrated into the chromosome (Colonna et al. 1995).

Another study analyzing H-NS structure and function in response to temperature concluded that H-NS tetramerization and activity is actually *higher* at elevated temperatures (Stella et al. 2006). Chimeric molecules were constructed whereby the N-terminal oligomerization and linker domains of H-NS were fused to the DNA-binding domains of phage repressors. Dimerization and oligomerization of these H-NS domains was determined through the use of reporter constructs situated downstream from different combinations of phage operator sequences. This experimental approach allowed the assay of β -galactosidase activity as a measure of DNA binding by the H-NS chimera as a dimer or higher-order oligomer (Stella et al. 2005). This approach, along with gel-filtration studies of purified H-NS protein, demonstrated that the formation of higher-order complexes is inhibited at temperatures below 25°C, providing a possible explanation for the anti-silencing of H-NS-regulated genes at lower temperatures. However, some *hns* mutant strains have a survival deficit at low temperatures (Dersch et al. 1994), and it is difficult to explain how H-NS could be inactive yet essential for survival under cold-shock. This model would predict the widespread activation of hundreds of H-NS-silenced genes during cold shock, which is in contrast to the approximately 25 genes actually induced under these conditions (La Teana et al. 1991).

H-NS has also been posited to act as a global regulator in response to osmolarity, and *hns* mutants display altered osmosensitivity (Barth et al. 1995; Hommais et al. 2001; Levinthal and Pownder 1996). H-NS was identified as a silencer of the osmoregulated *proU* operon and subsequently as a silencer of other osmoregulated genes including *osmC*, *osmY*, *otsAB* and the SPI-1 regulatory locus *hilA* (Atlung and Ingmer 1997; Olekhovich and Kadner 2006; Schechter et al. 2003). Force-extension measurements of individual λ DNA molecules complexed with H-NS indicate that H-NS increases the rigidity of DNA at temperatures below 32°C or in the presence of 200 mM KCl, which the authors interpreted to mean that H-NS alone might act as a

temperature or osmolarity sensor (Amit et al. 2003). However, these observations must be interpreted with caution as the concentrations of H-NS used in the study were in vast excess to the DNA which may lead to the “non-bridging” mode of DNA binding (Amit et al. 2004; Dame and Wuite 2003). Moreover, the use of KCl to alter osmolarity may have more pronounced effects on protein-DNA interactions than the physiologic salt potassium glutamate, whose cytoplasmic concentrations rise during osmolar stress (Gralla and Vargas 2006). The notion of H-NS as a simple osmosensor appears to be undermined by the same reasoning that questions its role as a thermosensor, namely that the expression of many H-NS silenced genes is unaffected by changes in osmolarity (Atlung and Ingmer 1997; Hulton et al. 1990). The *proU* promoter maintains much of its osmoregulation in the absence of H-NS (Fletcher and Csonka 1995), presumably due to H-NS independent changes in DNA supercoiling, and many other osmoregulated H-NS-silenced genes are directly activated by factors like OmpR or the alternative sigma factor σ^S . Furthermore many H-NS silenced *Shigella* virulence genes are inhibited by low osmolarity, even at temperatures normally associated with induction, which contrasts with another silenced locus, *hlyA* of *E. coli*, that is induced by low osmolarity (Carmona et al. 1993; Porter and Dorman 1994). Collectively, these observations illustrate the inadequacy of a simple model in which H-NS is unable to bind DNA under certain environmental conditions, but rather support a model in which H-NS and environmental parameters such as temperature and osmolarity have competing or interactive effects on DNA superhelicity and the expression of additional factors like sequence-specific DNA binding proteins that can act in concert to control the expression of individual loci. Specific examples where temperature or osmolarity appear to increase expression from certain promoters without altering the structure or oligomerization state of H-NS per se are discussed at the end of the following section.

13.9 Mechanisms of Anti-Silencing

Comparative analysis of several related bacterial genomes has revealed that many genes acquired via LGT are transient (van Passel et al. 2008). Most new sequences appear to be lost relatively quickly, never succeeding in establishing permanent residence in the genome either because they fail to enhance fitness or fail to properly integrate into the proper pre-existing regulatory circuits necessary for them to manifest their beneficial effects. With this in mind it is likely that most H-NS silenced genes, specifically those that are obtained by LGT, are eventually lost without ever having an appreciable effect on their host. This fact is underscored by the observation that despite large amounts of genetic exchange, any given isolate of *Salmonella* has stably maintained genes from only ~300 horizontal gene transfer events over the course of the past 100 MYr since the *Salmonellae* diverged from *E. coli* – a relatively slow rate of gene acquisition that averages one successful transfer event per 300,000 years (Groisman and Ochman 1997; Lawrence and Ochman 1998; McClelland et al. 2000).

In the rare cases throughout the evolution where H-NS silenced genes do find a useful function and enhance bacterial fitness it is important that the bacterial cell be able to activate such genes in a temporally appropriate manner in response to the correct set of environmental conditions. This poses a challenge given that the genes acquired via LGT do not necessarily encode their own regulatory proteins (although several xenogeneic elements do) nor do they necessarily have the necessary genetic information to properly integrate into the pre-existing regulatory networks of their new host cell. A survey of the literature reveals that many genes regulated by H-NS are also controlled by other global regulators including Lrp, Fis, IHF, CRP-cAMP, ppGpp and σ^S . This may be due in part to the fact that regulatory proteins with relatively degenerate target specificities (e.g. H-NS, HU, IHF, etc.) are more likely than highly-sequence specific transcriptional regulators to recognize newly introduced genetic elements and recruit them into pre-existing regulatory networks.

A comprehensive and well-written review on the various mechanisms by which xenogenic silencing can be overcome has recently been published (Stoebe et al. 2008). The authors note that antagonism of H-NS at various promoters appears to involve ad hoc solutions and that few specific mechanisms are widely employed. Rather, it appears that depending on the specific context a number of mechanisms can suffice to alter the H-NS nucleoprotein complex to enable transcription. This may be due in part to the fact that H-NS is a relatively poor DNA binding protein with fairly high off rates at most sites (Dame et al. 2006), the obvious exception being the recently identified high-affinity consensus site (Bouffartigues et al. 2007). Few anti-silencers act directly on H-NS to abrogate silencing, the notable exceptions being factors that antagonize H-NS by acting in a dominant-negative fashion (e.g. H-NST) and the H-NS antagonists encoded on the T7-family of phages (the 5.5 proteins). In most instances anti-silencing occurs through alteration of local nucleoid structure by protein protein binding or environmental influences, by displacement of H-NS through competition for DNA-binding sites, or a combination of these mechanisms. Several of these mechanisms are outlined in Fig. 13.5.

13.9.1 *Anti-Silencing – A Role for Promoter Strength?*

The *bgl* and *proU* promoters are each regulated by H-NS binding sites found upstream and downstream of the promoter; referred to as the upstream and downstream regulatory elements (URE and DREs) respectively (Dattananda et al. 1991; Overdier and Csonka 1992; Owen-Hughes et al. 1992; Schnetz 1995). At both promoters these elements appear to act cooperatively to silence transcription, perhaps by forming a repressive loop (Bouffartigues et al. 2007; Nagarajavel et al. 2007). In the absence of the URE the DRE elements of both promoters retain residual ability to downregulate transcription in an H-NS dependent manner and it is curious how such repression occurs in the absence of a second H-NS binding site (Nagarajavel et al. 2007). The ability of the *bgl* DRE to inhibit transcription may be, in part, due to hindering transcription elongation by RNAP, resulting in premature

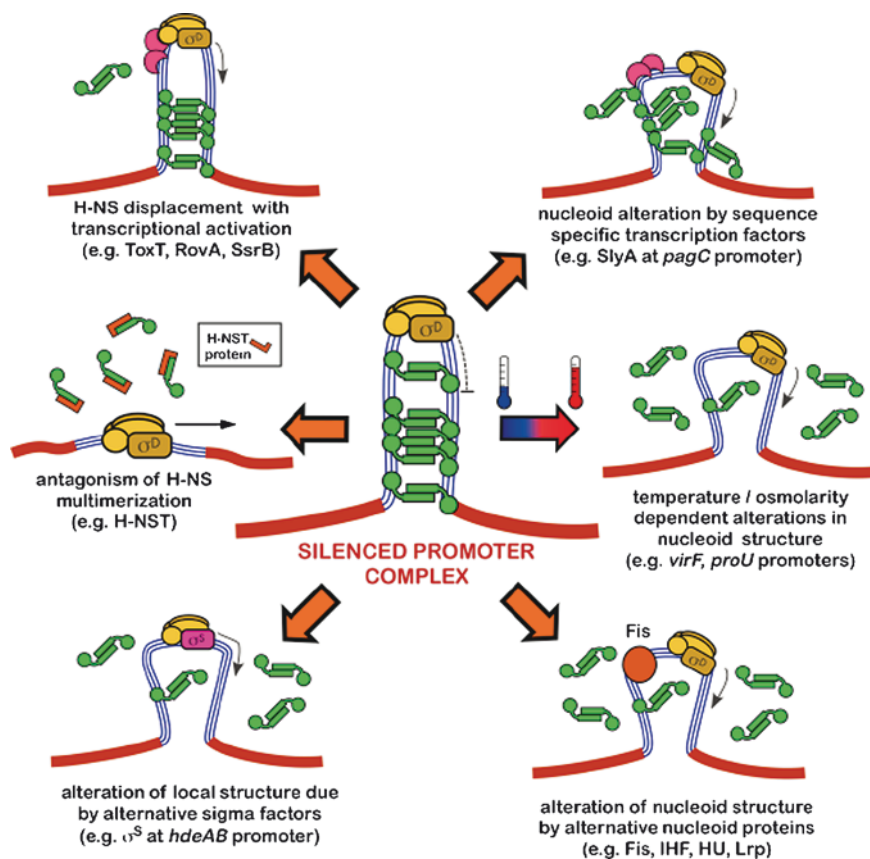


Fig. 13.5 Various mechanisms of anti-silencing. There are several distinct mechanisms of anti-silencing that have been elucidated from studies of different promoters, some of which are diagrammed here. Few of the mechanisms shown necessarily exclude others and it is quite often that more than one mechanism is involved in anti-silencing any given promoter. The *dps* promoter, for example, is regulated by a combination of nucleoid-associated proteins (*Fis* and IHF), changes in RNAP associated sigma factors, and sequence specific DNA binding proteins like *OxyR* (see text)

transcript termination by the transcription termination factor *Rho* (Dole et al. 2004b). However recent data suggests that downstream regulatory elements can block transcription at a step before open complex formation (Nagarajavel et al. 2007) and that this effect is most pronounced in the context of a weak promoter. Transcription from synthetic promoters containing the *bgl* and *proU* DRE is inhibited in conjunction with the weak P_{lacI} promoter but is mostly unaffected when transcription is driven by the relatively strong P_{tac} promoter. An intermediate strength promoter, P_{UV5} , displayed moderate inhibition by the *proU* and *bgl* DREs. This suggests that promoter strength may play a significant role in whether H-NS mediated silencing is effective for some sequences. This observation seems to be

important only for the downregulation mediated by the DRE in the absence of the URE. In combination, the *bgl* URE and DRE can effectively silence transcription directed by the *bgl*, *lac* and *lacUV5* promoters (Schnetz 1995). Although the relevance of promoter strength appears to be context dependent, it opens up a new line of inquiry into a previously underappreciated aspect regarding the mechanism of anti-silencing. Many transcription factors increase promoter strength and it is possible that this fact alone could account for their ability to serve as anti-silencers at some loci. Other ways that site-specific transcription factors contribute to anti-silencing, such as directly competing with H-NS for binding sites or altering nucleoid structure, are discussed below.

13.9.2 A Role for RNA Polymerase and Sigma Factors in Silencing and Anti-Silencing

As discussed briefly in Section 13.7.2, recent findings have suggested that RNA polymerase is not a passive player in H-NS-mediated silencing. Several findings have suggested that RNAP can act directly on the architecture of promoters to enhance silencing or anti-silencing by H-NS. In some cases, whether transcription occurs or not can depend on which sigma factor is associated with RNAP. Many loci silenced by H-NS are preferentially transcribed by RNAP in complex with the alternative sigma factor σ^S but cannot be transcribed by RNAP associated with the housekeeping sigma factor σ^D . Examples include *dps* (Grainger et al. 2008), *LEE1* (Ler) (Laaberki et al. 2006), *asr* (Seputiene et al. 2004), *gadBC* (De Biase et al. 1999; Waterman and Small 2003), *proU* (Rajkumari and Gowrishankar 2001), *hdeAB* (Shin et al. 2005), *csgBA* (Arnqvist et al. 1994; Olsen et al. 1993), *spvR* (Robbe-Saule et al. 1997), *csiD* (Marschall et al. 1998) and *yjiEFG* (Navarre et al. 2006). Although the basal promoter elements recognized by $\text{RNAP} \cdot \sigma^S$ and σ^D are highly similar, recent analyses of several σ^S -dependent promoters have revealed that they are less conserved and more AT-rich than $\text{RNAP} \cdot \sigma^D$ -dependent promoters (Becker and Hengge-Aronis 2001; Typas et al. 2007; Typas and Hengge 2006), features that might explain a greater tendency for $\text{RNAP} \cdot \sigma^S$ to initiate the transcription of newly acquired AT-rich genes. The finding that a stress-activated sigma factor can selectively drive the expression of certain H-NS-silenced genes (or, conversely, that H-NS can selectively silence $\text{RNAP} \cdot \sigma^D$) provides one pathway by which foreign genes can be rapidly assimilated into global regulatory networks. Anti-silencing by $\text{RNAP} \cdot \sigma^S$ might also account for the relationship between H-NS and regulation of loci in response to certain stresses (e.g., *proU* in response to osmolarity) without invoking a structural change in the H-NS protein itself.

Genetic experiments to determine the roles of H-NS and $\text{RNAP} \cdot \sigma^S$ at a given promoter are complicated by the fact that H-NS negatively regulates σ^S (Yamashino et al. 1995; Zhou and Gottesman 2006) and by the fact that H-NS and σ^S are inversely regulated by the small RNA DsrA, which enhances σ^S translation and stability while enhancing degradation of *hns* mRNA (Lease et al. 1998;

Majdalani et al. 1998). It should be noted that the effects of the DsrA RNA on H-NS levels were observed with DsrA in multicopy so the physiological relevance of this interaction remains unclear.

Biochemical studies, however, have generated some insight into how H-NS-mediated silencing can be selectively overcome by RNAP $\cdot \sigma^S$ at the *hdeAB* and *dps* promoters (Grainger et al. 2008; Shin et al. 2005). Analysis of the *hdeAB* promoter revealed that RNAP $\cdot \sigma^S$ is able to transcribe this H-NS-silenced locus much more efficiently than RNA polymerase associated with the housekeeping sigma subunit σ^D due to the different ways in which the DNA strand associates with the polymerase when complexed with the different sigma factors (Shin et al. 2005). Analysis of the role of RNAP $\cdot \sigma^S$ at the *dps* promoter suggests that the picture is slightly more complicated and involves additional factors including the nucleoid-associated protein Fis, as discussed in the following section (Grainger et al. 2008).

13.9.3 *Fis, HU and IHF*

Several loci have been identified where the abundant nucleoid-associated proteins Fis, IHF, and HU either augment or counter H-NS-mediated silencing. These findings are not particularly surprising given the fact that each of these proteins are highly pleiotropic regulators with >10,000 binding sites for each factor around the *E. coli* chromosome (Dame 2005; Grainger et al. 2007). The specific role played by these proteins at H-NS regulated promoters varies from promoter to promoter and no general mechanism applies to all situations. A genome wide analysis of binding sites for H-NS, Fis and IHF was recently performed by ChIP-on-chip and it was found that each of these proteins display biased targeting to intergenic regions (Grainger et al. 2007). The binding sites targeted by Fis and IHF were found to be AT-rich, confirming earlier observations. When binding was plotted against the AT-content of the microarray probe the curve was linear for Fis and IHF with increasing intensity at higher AT-values, but significant binding was still observed to probes of moderate GC-content. In contrast, H-NS was found to have a distinct cutoff value for AT-richness below which it fails to bind its target sequence (approximately 55% AT).

Fis is a small homodimeric DNA binding protein that binds to a moderately degenerate sequence motif that has been implicated as a global regulator in response to environmental conditions and growth rate (Finkel and Johnson 1992). Fis can regulate gene expression by binding in close proximity to a promoter but it can also influence gene expression by altering supercoiling or nucleoid structure in the vicinity of the gene (Travers et al. 2001). One mechanism by which Fis can affect local nucleoid structure is by inducing bends into DNA sequences of anywhere between 50° and 90° (Pan et al. 1996; Zhang et al. 2004). Expression of Fis is highly growth rate dependent where Fis levels are abundant during log phase growth and virtually absent in stationary phase in aerated cultures (Finkel and Johnson 1992). Fis however is easily detected in stationary phase under conditions

of limited aeration (Ó Cróinín and Dorman 2007). In the case of the *bgl*, *dps* and *virF* promoters Fis plays a negative regulatory role in conjunction with H-NS (Caramel and Schnetz 2000; Grainger et al. 2008; Prosseda et al. 2004), whereas Fis plays an activating role at the H-NS silenced *hns* and *rrnB* promoters (Falconi et al. 1996; Tippner et al. 1994).

The abundant nucleoid binding protein HU, like H-NS, is a dimer that binds DNA with highly degenerate sequence specificity. HU-like molecules are found in almost all eubacteria, usually existing as a dimer of two identical subunits. In *Salmonella* and *E. coli* HU exists predominantly as a heterodimer of HU α and HU β (encoded by the *hupA* and *hupB* genes, respectively) or sometimes as HU α and HU β homodimers. Arms of the HU dimer intercalate into the minor groove resulting in a severe bend in the target DNA substrate of up to 160°. Integration host factor (IHF) shares significant homology to HU and binds DNA with considerably more sequence specificity than HU does. Due to the radical changes in local nucleoid structure caused by the binding of HU and IHF it can be envisioned that depending on the particular spatial context of an HU/IHF binding site within a promoter HU/IHF binding could either enhance H-NS mediated silencing by bringing two distant H-NS binding sites in proximity or antagonize H-NS by spreading such sites apart. Scanning force microscopy supports a model where HU can antagonize H-NS through its ability to open up DNA, as opposed to the compacting effect that H-NS has on the nucleoid (Dame and Goosen 2002). Several examples have been cited where IHF exerts positive regulatory effects at H-NS silenced genes (Altuvia et al. 1994; van Ulsen et al. 1996). Accordingly, IHF has also been demonstrated to be a global activator of virulence gene expression in *Shigella*, *Salmonella* and *Vibrio* (Mangan et al. 2006; Porter and Dorman 1997; Stonehouse et al. 2008). Given their abundance and highly pleiotropic regulatory profiles it is not surprising that multiple nucleoid proteins and global regulators can cooperate to regulate certain promoters. For example Fis, IHF and H-NS all cooperate to either positively or negatively regulate transcription from the *dps*, *nir* and *csgD* promoters (Altuvia et al. 1994; Browning et al. 2000; Gerstel et al. 2003; Grainger et al. 2008).

Fis, HU and IHF can act as global regulators in response to environmental conditions in part because each of these proteins are themselves environmentally regulated. Fis, which is abundantly expressed in log phase and virtually absent in stationary phase, positively affects *rrnB* expression but is a negative regulator of Dps, a protein critical for virulence and oxidative stress (Ó Cróinín and Dorman 2007; Grainger et al. 2008; Tippner et al. 1994). Expression of Fis therefore correlates with the fact that rRNA expression is very high and Dps levels are very low in rapidly growing cultures, while the converse is true in stationary phase (Ali Azam et al. 1999; Grainger et al. 2008; Halsey et al. 2004). Furthermore Dps is critical for resistance to oxidative stress, which parallels the fact that Fis is virtually absent in stationary phase in shaking cultures but is expressed at considerably higher levels under microaerobic conditions where, presumably, the need for Dps is much lower (Halsey et al. 2004; Nair and Finkel 2004).

The *dps* promoter provides one example of the varied and interesting ways that pleiotropic regulators can cooperate to achieve precise control over gene expression

at a specific promoter in response to specific environmental conditions (Grainger et al. 2008). Dps expression is regulated by several pleiotropic factors including OxyR, RNAP $\cdot \sigma^S$, IHF, H-NS, and Fis depending on growth phase (Schnetz 2008). An elegant and extensive combination of high-throughput and biochemical assays has revealed that control of *dps* expression involves the selective repression of RNAP $\cdot \sigma^D$ by Fis during logarithmic growth and selectively granting access to RNAP $\cdot \sigma^S$ but not RNAP $\cdot \sigma^D$ during stationary phase growth (Grainger et al. 2006, 2008). Fis, when bound to a region within the core promoter of *dps*, can trap RNAP $\cdot \sigma^D$ at the promoter in a closed complex. Trapping of RNAP $\cdot \sigma^D$ at the promoter can also serve to exclude RNAP $\cdot \sigma^S$, although the physiological relevance of this blockage is unclear given that RNAP $\cdot \sigma^S$ is not the dominant form of polymerase during rapid growth. The Fis-directed block of RNAP $\cdot \sigma^D$ -mediated transcription can be overridden by the transcriptional activator OxyR in bacteria experiencing oxidative stress (Altuvia et al. 1994). In stationary-phase aerated cultures, when Fis is absent from the cell, H-NS can block access of RNAP $\cdot \sigma^D$ while enabling transcription from RNAP $\cdot \sigma^S$ perhaps by binding an element near the transcription start point or by a mechanism similar to that observed for the *hdeAB* promoter (Grainger et al. 2008). Stationary phase induction of *dps* is further enhanced by the action of IHF, although the exact role of IHF remains unclear (Altuvia et al. 1994).

13.9.4 Sequence Specific DNA Binding Proteins

A common way by which H-NS is antagonized at specific promoters is through the action of sequence-specific DNA binding proteins as was first shown for CfaD, an AraC-family transcription factor required for full expression of the CFA/I fimbrial operon in enterotoxigenic *E. coli* (Jordi et al. 1992). The mechanisms by which such “classical” transcriptional activators antagonize H-NS are unclear but are unlikely to involve direct interactions with H-NS. Evidence instead suggests that anti-silencing occurs either by displacing H-NS through competition for a common binding site or by invoking structural alterations upon DNA binding. Synthetic promoters have been constructed that allow anti-silencing even by non-native DNA-binding proteins such as TyrR or the Lac and λ repressors (Caramel and Schnetz 1998; Gowrishankar and Pittard 1998), arguing that any DNA-binding protein with sufficient affinity might be able to increase expression from an H-NS regulated promoter in the correct context. The recognition sites for VirB (Turner and Dorman 2007), ToxT (Hulbert and Taylor 2002) and SlyA/RovA (Stapleton et al. 2002) are very AT-rich, and the critical recognition motif for PapB is a run of three adjacent T/A bases (Xia et al. 1998), which may in part explain why these proteins were selectively co-opted to act as anti-silencers.

The most obvious and direct way that a sequence-specific transcription factor could antagonize H-NS-mediated silencing is through direct competition for a critical binding site resulting in H-NS displacement. Indeed the binding sites of many transcription factors overlap demonstrated H-NS binding sites including, SlyA/

RovA at the *hlyE*, *rovA*, *pagC*, *ugtL*, *inv*, and capsule V gene cluster promoters (Corbett et al. 2007; Ellison and Miller 2006; Heroven et al. 2004; Perez et al. 2008; Wyborn et al. 2004), ToxT at the *ctx* and *tcpA* promoters (Hulbert and Taylor 2002; Yu and DiRita 2002), HilC, HilD and RtsA at the *rtsA* promoter (Olekhovich and Kadner 2007), and VirB at the *icsB* promoter (Turner and Dorman 2007). Results from few cases where H-NS displacement has been tested directly in vitro have not been consistent and support the notion that different mechanisms are at play at different promoters. VirB binding to the *icsB* promoter does not significantly alter the amount of H-NS from the promoter (Turner and Dorman 2007) nor does SlyA displace H-NS from the *Salmonella pagC* or *ugtL* promoters (Perez et al. 2008). SlyA does, however, displace H-NS from the *hlyE* promoter (Lithgow et al. 2007) and its close homolog RovA can effectively compete with H-NS for binding at the *inv* promoter (Ellison and Miller 2006).

At the promoters where H-NS is not significantly displaced by site-specific transcription factors it is inferred that subtle changes in the structure of the nucleoprotein complex must occur that enable RNAP to bind or form productive elongation complexes. At several promoters the binding of SlyA and RovA appear to alter the H-NS-DNA complex to allow other transcription factors to bind and subsequently interact with RNAP (Heroven and Dersch 2006; Navarre et al. 2005; Perez et al. 2008). The need for two separate functions has been demonstrated for many genes under the control of SlyA in *Salmonella* (Navarre et al. 2005; Norte et al. 2003; Shi et al. 2004). Microarray analysis revealed that the vast majority of SlyA-activated genes in *Salmonella* also require the PhoP/PhoQ two-component regulatory system for expression (Navarre et al. 2005). A recent analysis of the functional roles of PhoP and SlyA at the *Salmonella ugtL* and *pagC* promoters revealed that in the absence of H-NS, SlyA is dispensable and PhoP is capable of interacting with RNAP to activate transcription. In the presence of H-NS, however, PhoP is unable to exert a significant positive regulatory effect without SlyA (Perez et al. 2008). That some promoters have a regulatory hierarchy, where one factor is an H-NS antagonist while the other factor interacts with RNAP directly, may allow finer control in regulation either through the use of feed forward loops or by enabling multiple inputs to regulate activity.

Another anti-silencing mechanism has been invoked to explain the mechanism by which LeuO counters H-NS-mediated silencing at its own promoter (Chen and Wu 2005). In this model LeuO acts as a molecular boundary marker that blocks the spread of an H-NS-DNA filament that would otherwise encroach upon the *leuO* promoter. This model is supported by the observation that LeuO-mediated anti-silencing of a synthetic promoter containing a strong H-NS binding site is most apparent if the LeuO binding site lies between the H-NS nucleation site and the core promoter elements. The observed LeuO dependent activation was enhanced when two such promoters were placed on the same plasmid, suggesting a bridging (looping) mechanism was critical to establishing the boundary. This idea is further supported by experiments demonstrating that the LeuO binding sites can be functionally replaced with binding sites for LacI, another protein capable of looping DNA duplexes. It is important to note that the observations used to form this roadblock

model are inferential and no direct demonstration of an H-NS filament being blocked from a promoter was performed (e.g. footprint analysis of H-NS-DNA complexes in the presence and/or absence of LeuO). Also, there are some aspects of the roadblock model that disagree with other aspects of what is known about H-NS. Recent data suggest that H-NS dimers each bind independently to DNA and the observed cooperative behavior is due to proximity of DNA strands, as opposed to filament formation due to direct lateral protein-protein interactions (Dame 2008; Dame et al. 2006). If true, then any molecule, such as LacI, that can bring two adjacent strands of DNA together might be expected to enhance H-NS binding and promote filament extension, rather than block it.

At some promoters the role of site-specific DNA binding proteins is limited to antagonizing H-NS and they appear to be completely dispensable for upregulating gene expression in the absence of H-NS (Perez et al. 2008; Westernmark et al. 2000; Wyborn et al. 2004). Other site-specific H-NS antagonists can activate transcription even in the absence of H-NS (i.e., a positive effect of the transcription factor is observed in *hns* mutants) (Jordi et al. 1992; Murphree et al. 1997; Tran et al. 2005; Walthers et al. 2007; Yu and DiRita 2002). This suggests that they may not only alter the local nucleoid structure to disrupt H-NS mediated silencing but may also play a role in recruiting RNAP to the promoter. RNAP binding affects nucleoid structure however, and promoter strength can also play a role in overcoming H-NS mediated antagonism. For this reason it remains unclear whether transcription factors that recruit RNAP antagonize H-NS independently of their ability to recruit RNAP or whether their ability to bind to and enhance the activity of RNAP is a requisite step for their apparent anti-H-NS function. To address the relative contributions of binding vs. interaction with H-NS it may be useful to employ mutant activator proteins that retain their DNA binding ability but lack the ability to interact with RNAP (or vice versa with mutant RNAP). These types of experiments have been performed to analyze the roles RovA and RNAP each play at the *invA* promoter as well as the roles of IHF and RNAP at the phage Mu Pe promoter (Tran et al. 2005; van Ulsen et al. 1997a,b).

13.9.5 *Ler*

Ler, the *LEE* encoded regulator, is an unusual example of an H-NS antagonist because it is itself an H-NS family member – the only one described thus far to have the ability to activate transcription. Among the H-NS like molecules *Ler* has most similarity to the *Bordetella* BpH3 molecule (Elliott et al. 2000). *Ler* is encoded on the locus of enterocyte effacement pathogenicity island (LEE) of enteropathogenic and enterohemorrhagic *E. coli* (EPEC and EHEC, respectively) and is responsible for activation of genes contained within the island in response to certain environmental conditions. The LEE island encodes a type-III secretion system (T3SS) that is largely responsible for the attaching and effacing (AE) intestinal lesions caused by EPEC and EHEC (Caron et al. 2006; Kaper et al. 2004; Nougayrede et al. 2003).

It is composed of several operons designated as *LEE1* to *LEE5*, *espG*, and *grlRA*, as well as the EPEC specific genes *map* and *escD* (Mellies et al. 2007).

The regulation of the entire LEE island centers on control of the *LEE1* operon, which encodes Ler. *LEE1* is negatively regulated by H-NS as well as by Ler itself, presumably in a negative feedback loop to control its own expression levels (Berdichevsky et al. 2005). H-NS-mediated silencing of *LEE1* occurs at 27°C and is anti-silenced at 37°C (Umanski et al. 2002). In addition to temperature, a large number of other environmental inputs play a role in the regulation of LEE, including quorum sensing, pH, calcium, iron, and ammonium. Accordingly, a similarly large number of proteins including Fis, IHF, GadX, BipA, GrlRA, PerC, and two phage-encoded proteins called Pch also play a role in controlling LEE expression, primarily through their effects on *LEE1* (Abe et al. 2008).

H-NS also negatively regulates several other LEE operons including *LEE2*, *LEE3* and *LEE5* and *grl* and it is at these operons where Ler plays a positive role in gene expression by antagonizing H-NS mediated silencing (Barba et al. 2005; Bustamante et al. 2001; Haack et al. 2003; Sperandio et al. 2000; Umanski et al. 2002). Ler also positively regulates gene expression by antagonizing H-NS at chromosomal loci encoded outside LEE including the long polar fimbrial (*lpf*) locus (Torres et al. 2007). A recent ChIP-on-chip study of Ler binding sites in *E. coli* O157:H7 strain Sakai found 59 binding sites throughout the *E. coli* chromosome that distributed largely to the Sakai-specific genomic islands (S-loops) (Abe et al. 2008). The mechanism by which Ler antagonizes H-NS mediated silencing remains unclear as do the specific factors that play a role in determining whether Ler acts positively (e.g. *LEE2*, *LEE3*, *LEE5* or negatively (e.g. at *LEE1*) on gene expression.

Recently, the group of Mellies has examined the functional domains of Ler that are necessary for transcriptional activation of the *LEE5* promoter (Mellies et al. 2009). Ler has a significantly truncated N-terminal domain compared to H-NS that is predicted to lack the first two alpha helices and has an 11 amino acid extension at the C-terminal DNA binding domain. Mutagenesis and domain swap experiments where the N-terminal, C-terminal, and central domains were exchanged between Ler and H-NS were performed to determine if any of these structural differences are responsible for the phenotypic differences between the two molecules. Surprisingly, the 11 residue C-terminal extension was dispensable for Ler function and the N-terminal domain of Ler could be exchanged with its H-NS counterpart without significantly altering the ability of the molecule to activate expression at *LEE5*. The N-terminal domain of Ler is not dispensable, however, as mutations in this domain that abolish the ability of Ler to form dimers also abolish its activity to activate gene expression (Sperandio et al. 2000). Swapping either the C-terminal or linker domain with that of H-NS led to a strong reduction in the ability of the hybrid molecule to stimulate *LEE5* expression. Further mutagenesis of the Ler linker and DNA-binding domains found these regions were essential for Ler activity (Mellies et al. 2008).

Together these studies indicate that Ler activity requires dimerization but that the exact sequence of the N-terminal dimerization domain is of lesser importance.

The DNA binding properties of Ler are probably sequence-specific and differ from that of H-NS since its C-terminal domain cannot be swapped with that of H-NS. The domain swapping study did not test whether the Ler constructs with mutations in the linker domain retained their DNA-binding activity and footprinting and competition assays have yet to be carried out with these chimeras, leaving the exact role of the C-terminal and linker domains unclear. If mutations in Ler can be identified that retain DNA binding activity but fail to antagonize H-NS it may point to a novel mechanism by which H-NS can be antagonized.

13.9.6 H-NST and Dominant Negative H-NS molecules

H-NS N-terminal domains, when expressed alone without a C-terminal DNA-binding domain, have been shown to act in a dominant negative fashion, inhibiting H-NS mediated silencing presumably by interfering with the ability of H-NS to dimerize or form higher-order complexes (Banos et al. 2008; Ueguchi et al. 1996; Williams et al. 1996). It was therefore interesting when ORFs encoding proteins with similarity to H-NS N-terminal domains were identified in pathogenicity islands of some pathogenic *E. coli* strains (Williamson and Free 2005). These molecules have been designated H-NST, for H-NS “truncated”. The H-NST molecule encoded on the *serU* island of EPEC (H-NST_{EPEC}) could co-purify with endogenous H-NS and its expression led to significant upregulation of the *proU* and *bgl* operons, indicating that H-NST_{EPEC} is a *bone fide* H-NS antagonist. A similar H-NST from UPEC (H-NST_{UPEC}) was found to have much weaker activity and this was traced to a change in a single residue that diminished the ability of the protein to interact with H-NS. A recent study has found that H-NST_{EPEC} has the ability to also antagonize H-NS when expressed in *Yersinia* (Banos et al. 2008).

The existence of the H-NST like molecules as stable components of pathogenicity islands is somewhat surprising since their expression would be predicted to cause genome-wide dysregulation at several loci that would presumably lead to a fitness defect akin to what is observed in an H-NS mutant. Nevertheless, the fact that these proteins are expressed during log phase growth under laboratory conditions and the fact that H-NST molecules have persisted on more than one island over what appears to be a relatively long span of evolutionary time suggests that they indeed play a functional role in regulation.

13.9.7 Phage T7 5.5 Protein

In 1993, Liu and Richardson serendipitously found that 5.5, a small 11 kDa protein from coliphage T7 strongly associated with H-NS during their attempts to purify the phage protein (Liu and Richardson 1993). Expression of the 5.5 protein in *E. coli* leads to increased transcription at the *proU* locus and although no mechanism

has been formally described for its activating function it appears to be necessary for full phage virulence. A mutant in 5.5 that abrogated the ability of T7 to grow on lambda lysogens abolished its interaction with H-NS, suggesting that the *rbl* (restricted by lambda) phenotype in this mutant may have something to do with its ability to antagonize H-NS. The genome of T7 is AT-rich, particularly in the promoters of the late genes that encode the head and tail proteins. It is tempting to speculate that a primary role of H-NS is to block phage replication via xenogeneic silencing and that the phage employs 5.5 to counteract this ability of H-NS. It is equally likely, however, that the phage simply uses H-NS for its own ends, perhaps to regulate its late promoters in order to coordinate timing of late protein synthesis to prevent premature synthesis of phage capsid proteins. This hypothesis is consistent with the fact that 5.5 is encoded in the “middle” (as opposed to early) operon and the fact that the phage “late” promoters are very AT-rich. Homologues of the 5.5 protein are found in the subset of T3/T7 phages that infect enteric bacteria but have been identified in no other phage groups sequenced thus far leaving it an open question whether other AT-rich phages employ mechanisms to counteract silencing by H-NS.

13.9.8 Temperature, Osmolarity and Their Effects on Promoter Activity and H-NS-Mediated Downregulation

Previously it was argued that H-NS does not respond in a simple and universal manner to environmental factors such as osmolarity and temperature. It is clear, however, that changes in these conditions play a direct and important role in counteracting or augmenting H-NS-mediated silencing at specific promoters. It has long been known that the geometric structure of DNA can be altered in response to conditions such as hydration, supercoiling, temperature, and osmolarity and that the specific changes that occur are sequence dependent. Given the apparent sensitivity of H-NS to changes in nucleoid structure that are induced by protein binding it may also be expected that changes in nucleoid structure in response to environmental conditions may also alter H-NS function at some promoters.

For many loci it remains unclear whether changes in gene expression are due to direct environmentally-induced perturbations in promoter structure or whether the response is due to upstream regulators that themselves respond to the specific environmental condition. Studies on the regulation of the H-NS regulated invasin (*invA*) gene of *Yersinia* revealed that under laboratory conditions invasin expression is maximal at 25°C and downregulated at 37°C. The cause of this thermoregulation was determined to be independent of temperature-mediated effects at the *invA* promoter, but rather due to the increased expression of the H-NS antagonist RovA at room temperature (Heroven et al. 2004). Expression of RovA is itself H-NS regulated which may indicate that temperature induced changes in H-NS-mediated silencing at the *rovA* promoter are involved in controlling a regulatory cascade for

virulence gene expression (Heroven and Dersch 2006; Heroven et al. 2004). A similar regulatory cascade could be in place at the *Salmonella* *ssrB* gene, which controls the expression of numerous H-NS regulated virulence genes including the SPI-2 type-III secretion system. SsrB, like RovA, is an activator of H-NS silenced genes that is itself negatively regulated by H-NS (Walthers et al. 2007). Recent work has shown that SsrB expression is temperature dependent and is downregulated at room temperature in a manner that requires H-NS although the mechanism by which this occurs remains unexplored (Duong et al. 2007).

One case where temperature has been implicated as playing a direct role in altering H-NS-mediated silencing is at the *virF* locus encoded on the *Shigella* virulence plasmid (Prosseda et al. 1998, 2004). VirF positively regulates the structural gene *icsA* as well as another major plasmid-encoded virulence regulator in *Shigella*, VirB, each of which are themselves silenced by H-NS (Beloin and Dorman 2003). VirF is the first positive activator in the regulatory hierarchy that controls virulence gene expression in *Shigella*. The *virF* promoter is antagonized by Fis and H-NS through cooperative interactions with an intrinsically curved segment of DNA (Falconi et al. 2001; Prosseda et al. 2004). At room temperature the intrinsic bend and Fis cooperate to align two H-NS binding sites in an orientation that facilitates the formation of a repressive nucleoprotein complex. At elevated temperature the bend center shifts dramatically such that the two H-NS binding sites are misaligned and a repressive complex cannot be maintained.

The fact that temperature can directly affect the expression of a single H-NS-regulated gene product (e.g. RovA, VirF, and possibly SsrB), which in turn can antagonize H-NS at dozens of other genes, points out that it is important not to over-interpret the findings that many H-NS regulated genes are responsive to temperature. The response of any given H-NS-silenced locus to changes in temperature (e.g. *invA*, *virB*, *SPI-2*) is just as likely to be indirect. Unfortunately, determining if a given promoter is directly responsive to temperature is difficult and requires an intimate knowledge of the promoter structure and the complete set of regulatory factors to recapitulate temperature mediated regulation in vitro. A plausible model for temperature dependent regulation of the *virF* promoter came about only through extensive genetic and biochemical analysis. Even in cases where a change in H-NS affinity due to temperature can be demonstrated in vitro does not necessarily mean that the specific promoter will be temperature dependent in vivo (Badaut et al. 2002; Bouffartigues et al. 2007; Stoebel et al. 2008).

In summary, the above examples serve to highlight the fact that it is currently impossible to infer the mechanism by which any given H-NS antagonist works at one promoter given what is known about its function at another promoter. The degree and diversity of mechanisms by which H-NS-mediated silencing is antagonized argues against the existence of any “universal mechanism” for either silencing or anti-silencing. Further, the fact that no two promoters appear to be regulated in the same way underscores the *ad hoc* nature by which H-NS regulated genes integrate into their appropriate cellular regulatory networks during the evolutionary process (Stoebel et al. 2008).

13.10 Other Effects of H-NS on Lateral Gene Transfer and Mobile DNA

H-NS can also regulate mobile and xenogeneic DNA in contexts beyond its ability to silence expression of AT-rich genes. In numerous studies H-NS has been demonstrated to play a role in conjugation, recombination and transposition but there are no simple rules regarding its exact function in each process and H-NS-dependent phenotypes can differ dramatically even between seemingly closely related systems. This will be illustrated by the examples discussed below.

13.10.1 H-NS and Conjugation

Given its predilection to regulate genes obtained via LGT it may not be surprising that H-NS has been observed to play a role in regulating conjugal transfer of episomes. The transfer apparatus of the *E. coli* F-plasmid is under the control of at least three plasmid-encoded regulatory proteins: TraM, TraJ and TraY. In wild-type cells F-plasmid conjugation is strongly inhibited in stationary phase, in part due to H-NS. The promoters of *traJ*, *traM* and *traY* are all bound and silenced by H-NS and the regions bound by H-NS at the *traJ* promoter have been determined to be both curved and very AT-rich by DNAase I footprinting analysis (Will and Frost 2006; Will et al. 2004). Accordingly, F-plasmid transfer can occur well into stationary phase in *hns* mutants.

Another study on the F-like plasmid, pRK100 came to strikingly opposite conclusions about the role of H-NS and conjugal transfer (Starcic-Erjavec et al. 2003). H-NS binds the *traJ* promoter but, surprisingly, *traJ* expression is lower and plasmid transfer is decreased over 500-fold in *hns* mutants. This data indicates that H-NS plays a positive regulatory role in conjugal transfer of pRK100 but the mechanism by which it does is entirely unclear. It should be noted that mutations in Lrp also decreased expression of the *traJ* gene but did not similarly affect the level of conjugation, suggesting that the reduction of *traJ* synthesis in *hns* mutants does not entirely account for the observed defect in transfer. The data regarding H-NS and its role in pRK100 conjugation are unusual in many regards including a rare example of positive regulation. The findings also indicate that the role H-NS plays with regard to the F-plasmid cannot necessarily be extended to other closely related similar plasmids.

The conjugative IncH1 plasmid, R27, encodes proteins that are homologous to H-NS (ORF164) and Hha (ORF182) (Forns et al. 2005). Deletion of either one or both of these plasmid-encoded genes leads to an increase in conjugation frequency indicating that the plasmid encoded molecules are functional and play a negative regulatory role. Additional mutations in the chromosomal copies of either H-NS or Hha further increased conjugation suggesting significant functional redundancy. This redundancy is further supported by the fact that the H-NS and Hha-like proteins

encoded on R27 could partially complement mutations in the chromosomal copies of *hns* and *hha* (Forns et al. 2005).

13.10.2 Positive and Negative Effects of H-NS on Transposition

H-NS has been found to be a host-factor that alters the activity or specificity of a number of transposable elements and insertion sequences. In the case of phage Mu H-NS appears to silence transposition whereas in other cases (IS1 and Tn10) it appears to be a necessary co-factor for the transposition reaction. A positive role for Hha in transposition has also been observed but not extensively characterized (Balsalobre et al. 1996; Mikulskis and Cornelis 1994). The mechanisms by which H-NS alters transposition differ with each specific example, suggesting that no simple rule can be applied when discussing the role of H-NS and its relationship to transposons.

13.10.2.1 Positive Effects of H-NS on Transposition

H-NS can facilitate transposition for at least four transposable elements: IS1, Tn5, Tn10, IS903 and Tn552 (Shiga et al. 2001; Swingle et al. 2004; Whitfield et al. 2009). The mechanistic details by which H-NS can facilitate transposition have been worked out in the most detail by David Haniford and colleagues studying the role of H-NS in Tn10 transposition. Over-expression of H-NS increases transposition approximately three-fold in papillation assays while transposition is severely impaired by overexpression of a P116S mutant derivative of H-NS or in *hns* mutants (Ward et al. 2007). Transposase binding to the terminal inverted repeats during formation of the initial transpososome leads to distortions in flanking donor DNA that may provide a high affinity site for H-NS binding. Subsequent interactions with transposase enzyme enable H-NS to bind DNA within the core of the transpososome (Ward et al. 2007). When bound in this way H-NS appears to play a role in maintaining the transpososome in an unfolded state, a conformation that promotes inter-molecular transposition reactions as opposed to self-destructive intramolecular transposition events (Singh et al. 2008; Wardle et al. 2005). H-NS has also recently been demonstrated to play a role in the correct folding and assembly of the Tn5 transpososome in vitro and is necessary for efficient transposition in vivo (Whitfield et al. 2009)

Tn5 and Tn10 serve as interesting examples of how a “selfish” mobile genetic element can co-opt a host molecule for its own ends. Perhaps H-NS was selected for this function due to its ability to bind certain structural features in DNA, notably the distortions that may occur at the transposon ends. It is also possible that its role in transposition is not only structural but also regulatory. Some transposons may gain the ability to sense certain conditions that are favorable for transposition by integrating into H-NS-dependent regulatory circuits (Haniford 2006). Given that most stresses that

increase transposition are thought to decrease binding of H-NS (i.e. lead to activation of H-NS silenced genes) it is not exactly clear how such regulation would occur. Nevertheless, either hypothesis is intriguing and merits further exploration.

Another transposable element where the role of H-NS has been examined in some detail is *IS1*. Overexpression of the multisubunit *IS1* transposase in wild-type *E. coli* leads to induction of the SOS response (Lane et al. 1994), a phenotype that is abolished in *hns* mutants (Rouquette et al. 2004; Shiga et al. 2001). Mutations in *hns* also lead to a 100-fold decrease in *IS1* transposition. The defect appears to be primarily due to instability of the *IS1* transposase in H-NS deficient cells as opposed to H-NS binding the transposon directly (Rouquette et al. 2004). Neither the C-terminal nor N-terminal domains of H-NS are necessary for *IS1* transposition, which implicates the central flexible linker as playing a critical role for *IS1* transposase stabilization (Shiga et al. 2001). The observed *IS1* transposase instability, and the effects of an *hns* mutation, can be reversed in cells lacking the Lon protease and partially reversed in cells lacking the *ssrA* gene, encoding a factor that degrades mistranslated proteins (not to be confused with the SPI-2 encoded virulence regulator in *Salmonella* of the same name) (Rouquette et al. 2004).

13.10.2.2 Countering Transposition – Bacteriophage Mu

Bacteriophage Mu behaves like a transposon both when it integrates into the host genome and during the replicative (lytic) stage of its lifecycle. Mutations in *hns* were observed to lead to an increase in transposition of a Mu prophage defective for DNA packaging and lysis but still competent for transposition and replication (Falconi et al. 1991). The repressive effect H-NS has on Mu transposition may be due to its ability to repress the Mu early promoter (Pe) as observed both in vivo and in vitro (Kano et al. 1993; van Ulsen et al. 1996). H-NS binds a large segment of the Pe promoter and H-NS mediated repression is, in part, antagonized by IHF (van Ulsen et al. 1996). Therefore, in the case of bacteriophage Mu H-NS represses the lytic cycle not by acting on transposition directly, as it does for *Tn10*, but rather by silencing the expression of a set of genes required for the early stage of the Mu lytic cycle.

13.10.2.3 A Role for H-NS in Transposon Target Site Selection?

There is some compelling evidence that H-NS can play an important role in biasing the sites that transposons are targeted to. A recent genome-wide analysis of *IS903* targeting revealed that mutations in *hns* not only decreased the total amount of transposition, but also dramatically altered the locations into which *IS903* was targeted (Swingle et al. 2004). *IS903* had an increased preference for intergenic regions on the chromosome in *hns* mutants, indicating that perhaps H-NS was altering the nucleoid structure of these regions in wild-type cells to make them less permissive for transposition. This study also produced some indirect evidence that target

site selection for *Tn10* was also altered in *hns* mutants. Interestingly *Mu* and *IS1* each have separate well-defined insertion hot-spots immediately upstream of the *bgl* operon that disrupt a strong H-NS binding site and lead to the activation of the cryptic *bgl* locus, suggesting that H-NS dependent effects on nucleoid structure at this locus may create this hotspot (Manna et al. 2001).

13.10.3 *H-NS and Site-Specific Recombination: Phase Variable Switching of the Type I Fimbriae*

The type I fimbriae (pili) are an important virulence factor of many pathogenic strains of *E. coli* (Connell et al. 1996). Fimbrial expression is phase variable due to its control by a promoter region (*fimS*) that can switch between “on” and “off” orientations. This switching process involves a site-specific recombination reaction mediated by any one of a few different integrase family recombinases (Bryan et al. 2006; Klemm 1986). The two primary *fim* recombinases, FimB and FimE have different effects on *fimS*. FimB mediates either the “on-to-off” or “off-to-on” switching of *fimS* whereas FimE shows a strong bias toward switching *fimS* from “on-to-off”, but not the reverse (Klemm 1986). Under normal laboratory growth conditions in wild-type cells the activity of the FimE recombinase is dominant but the degree of bias can change in response to environmental factors including temperature (Gally et al. 1993).

It has long been known that H-NS plays a role in regulating the inversion of *fimS*, along with the nucleoid-binding proteins IHF and Lrp (Gally et al. 1993). Null mutant *hns* strains display elevated levels of switching, up to 100-fold higher than that observed for wild-type cells, rapidly reaching a steady state population that is almost equally balanced between cells in the “off” state and cells in the “on” state (Higgins et al. 1988; Kawula and Orndorff 1991; Spears et al. 1986). The effect of H-NS on *fimS* switching appears to be due to multiple factors including the fact that FimB expression is silenced by H-NS (Donato and Kawula 1999; Donato et al. 1997). Furthermore H-NS may make a nucleoprotein complex at *fimS* that biases switching toward the “off” state and the effects of this switch (O’Gara and Dorman 2000). Evidence suggests that the ability of H-NS to affect *fimS* directly depends in part on active transcription within and adjacent to the *fimS* locus, implying that local nucleoid structure (i.e. transient changes in accessibility or supercoiling due to transcription) mediates H-NS binding or function at this locus.

In the enteric bacteria the expression of fimbriae and flagella almost never occur simultaneously as the role of each are mutually incompatible (“stick vs. swim”). In that regard it is notable that H-NS is a primary negative regulator of fimbrial expression but acts in a positive fashion on the expression of flagella. At least part of the H-NS-dependent reciprocal regulation of flagella and fimbriae is due to the fact that a fimbrial gene product PapX represses the flagellar genes by antagonizing the *flhDC* promoter (Lane et al. 2007). Derepression of fimbrial *papX* like genes may explain why *hns* mutants are deficient for motility.

13.11 What We Can Learn from H-NS Phylogeny

Unlike the HU proteins and the SMC complex proteins (eg. MukBEF in *E. coli*), which are widespread in the eubacteria, the H-NS like molecules are not widely distributed among bacterial species and identifiable homologues are found only in certain members of the α -, β - and γ -proteobacteria (Tendeng and Bertin 2003). Furthermore the diversity in primary sequence among the various H-NS family members can be unusually high when compared to the relatedness of the species. For example the two H-NS-like molecules of *Pseudomonas aeruginosa*, MvaT and MvaU, share only ~18% identity with *E. coli* H-NS despite the fact that both bacteria are γ -proteobacteria (Tendeng et al. 2003b). In contrast, Fis, HU, DNA gyrase, RNAP subunits, and TCA enzymes of *Pseudomonas aeruginosa* PAO1 and *E. coli* strain K12 all share approximately 60% identity between the two species (Blattner et al. 1997; Riley et al. 2006; Stover et al. 2000). The reasons underlying the relatively radical divergence between H-NS and MvaT/MvaU are unclear but these findings suggest that the H-NS like molecules may be under strong selective pressure to change. Interestingly the highest level of conservation is found in the C-terminal DNA binding domain, presumably because this region is constrained to maintain its DNA binding function while, in comparison, multimerization domains have considerably greater ability to alter sequence without disrupting function.

It is likely, although unproven, that H-NS plays a role in defense against phages and transposons and this may place a strong selective pressure on this family of molecules to diverge in their primary sequence. Most phage genomes are relatively AT-rich compared to their respective hosts and, as discussed in an earlier section, several T7 phages encode “5.5 proteins” that serve as H-NS antagonists (Liu and Richardson 1993; Rocha and Danchin 2002). Examination of the T7-like phages reveals no 5.5 close homologue is present in any of the T7 phages that infect *Pseudomonas*. It is unclear if these phages have modified their antagonist to counteract silencing by MvaT/MvaU to such a degree that it cannot be identified through homology, or if they simply have no functional equivalent to this protein.

In addition to the relatively high sequence variability found among H-NS family members, the number of H-NS-like molecules harbored by different proteobacterial species can vary widely from 1 for *Yersinia* sp. and *Haemophilus* sp. to eighteen for *Burkholderia vietnamiensis* strain G4 (Tendeng and Bertin 2003). The reasons why some species harbor an abundance of H-NS like molecules are unclear. However most species only harbor one to three H-NS paralogues that are common to all members (e.g. H-NS and StpA for *E. coli*) with the remainder encoded on or near mobile genetic elements and pathogenicity islands (e.g. Sfh, Ler, H-NST). *B. vietnamiensis* strain G4, harbors three chromosomes and five smaller plasmids. Two of the chromosomes (1 and 2) are highly homologous to the chromosomes of all *Burkholderia* sp. while the third chromosome is considered a “mega-plasmid” with a highly mosaic structure that shares almost no homology to any other chromosome found in bacteria, including other members of *Burkholderia*. Chromosomes 1 and

2 encode three H-NS molecules that are highly conserved and universally present in all *Burkholderia*. In contrast, the third chromosome encodes nine H-NS-like proteins while the remaining six are distributed among the smaller plasmids. Given their high degree of conservation it is likely that the three H-NS-like proteins encoded on chromosomes 1 and 2 serve a central role in gene regulation or chromosomal compaction. The functions of the other fifteen H-NS homologs, which may be only transient residents of the strain G4 genome, can only be speculated at this point.

Although much can be learned through the analysis of species that encode many H-NS homologs it is also informative to note which species have lost H-NS. The only members of the enteric bacterial species that do not encode an H-NS like molecule are the endosymbionts *Buchnera aphidicola* and *Blochmania floridanus* species that diverged from their common ancestor with *E. coli* approximately 50–100 MYr ago by establishing residence in the cells of aphids and carpenter ants, respectively (Gil et al. 2003; Tamas et al. 2002). These endosymbionts diverged from a common ancestor to the enteric lineage at approximately the same time as *E. coli*, *Yersinia* and *Salmonella* each arose (approximately 50–150 MYr ago). Because of their restricted niche *Buchnera* and *Blochmania* have small genomes that, like the genomes of all endosymbionts, are also markedly AT-rich (~25% GC). Due to their long-term sequestration inside of a eukaryotic host cell *Buchnera* and *Blochmania* have lost all contact with the outside microbial world. One possible interpretation of these combined observations is that in the absence of a “social” lifestyle, H-NS can become dispensable. Furthermore it suggests that maintaining an elevated genomic GC-content, like many free-living bacteria do, may be an investment that is made to facilitate the recognition of foreign sequences (Navarre et al. 2006).

13.12 Functions of H-NS-Like Proteins

With the exception of StpA there is very little known about the function of other H-NS-family members. The few paralogs that have been examined to date have been shown to functionally complement phenotypes such as motility, serine sensitivity, or salicin utilization when ectopically expressed in *E. coli hns* mutants but virtually nothing is known about the role of most of these proteins in their natural context (Tendeng et al. 2000, 2003a,b; Tendeng and Bertin 2003). Such screens are informative and have provided important information however some phenotypes of *hns* mutants can be complemented with proteins that likely have nothing to do with nucleoid structuring or xenogeneic silencing including the kin17/btcd, a zinc finger DNA-binding protein in mammals that has no significant primary amino acid homology to H-NS yet can complement an *hns* mutation in *E. coli* (Timchenko et al. 1996; Tissier et al. 1996). The kin17 protein is stress-activated and appears to be involved in DNA replication in mammalian cells (Masson et al. 2003). Whether kin17 has any functional equivalence to H-NS in its native context has not been addressed.

Progress has been made in our understanding of the functions of a handful of H-NS family members in their natural context and a few important examples are highlighted below (with the exception of Ler, which was discussed in Section 13.9.5).

13.12.1 StpA – An H-NS Parologue with a Unique Function?

The genomes of several enteric bacteria including *Salmonella*, *Escherichia* and *Shigella* (but not *Yersinia*) encode a second H-NS parologue, StpA, with ~58% identity to H-NS and very similar domain structure (Cusick and Belfort 1998; Williams et al. 1996; Zhang and Belfort 1992). Like H-NS, StpA is capable of constraining supercoils and bridging DNA with a preference for AT-rich sequences with planar curvature (Ali Azam et al. 1999; Dame et al. 2000, 2005; Zhang et al. 1996). StpA is a competent DNA binding protein that has four- to six-fold greater DNA binding affinity than H-NS for at least one prototypical H-NS target (Sonnenfield et al. 2001). Structural similarity between StpA and H-NS is supported by the fact that antibodies against StpA can also bind H-NS (Ali Azam et al. 1999; Sonnenfield et al. 2001) and the fact that StpA is capable of interacting with H-NS through its N-terminal dimerization domain (Cusick and Belfort 1998; Deighan et al. 2003; Johansson et al. 2001; Williams et al. 1996). StpA is stable in wild-type cells but is rapidly degraded by Lon protease in the absence of H-NS, an effect that is exacerbated further in the absence of YdgT or Hha (Johansson and Uhlin 1999; Paytubi et al. 2004). Its stability can be greatly enhanced by a specific mutation within the StpA N-terminal domain (F21C) (Johansson et al. 2001). StpA is also stabilized in *hns* mutants when co-expressed with the N-terminal dimerization domain of H-NS containing residues 39–60. Together these facts suggest that the ability of StpA to form heterodimers with H-NS is physiologically relevant and occurs in the cell under normal conditions (Johansson et al. 2001; Johansson and Uhlin 1999).

ChIP-chip analysis of StpA binding throughout the *E. coli* chromosome has been performed both in the presence or absence of H-NS (Uyar et al. 2009). In wild type *E. coli* cells the binding profiles of epitope-tagged H-NS and StpA were virtually superimposable and deletion of *stpA* resulted in no significant change in the binding profile of H-NS. However, deletion of *hns* resulted in a severe reduction in the number of regions bound by StpA, with approximately two-thirds of the StpA-bound regions requiring H-NS for binding. It remains unclear why StpA binding was altered at specific locations in the absence of H-NS, however heterodimer recruitment is one possibility (i.e. that StpA association with certain regions only occurs with H-NS/StpA heterodimers). Another possibility is that local nucleoid changes induced by the absence of H-NS caused a decrease in the affinity of StpA for certain regions of the chromosome. The observed binding differences, however, do not appear to be due to StpA stability as the loss of binding for specific sites in *hns* mutants was observed even for the more stable F21C StpA mutant.

The function of StpA remains unclear since few, if any, phenotypes have been attributable to *stpA* mutants that contain a wild-type copy of *hns* (Bertin et al. 2001;

Zhang et al. 1996). This is due in part to the fact that expression of StpA mRNA is fairly low in wild-type cells under most conditions, but its expression is enhanced approximately 20-fold in H-NS mutants (Sonnenfield et al. 2001). As might be expected given the similar binding profiles of H-NS and StpA, StpA overexpression from plasmid vectors can complement several phenotypes of *hns* mutants including silencing arginine decarboxylase, *proU*, *bgl*, the synthetic 5A6A *galP1* promoter, and *hns* (Johansson et al. 2001; Shi and Bennett 1994; Williams et al. 1996; Zhang et al. 1996), but under the control of its native promoter it is only capable of compensating for a subset of H-NS dependent phenotypes (Free and Dorman 1997; Wolf et al. 2006). Because of its instability in the absence of H-NS, StpA protein does not accumulate in *hns* mutants beyond 10% of the levels observed for H-NS in wild-type cells, which may partly account for the mild effects observed in *stpA* mutants (Sonnenfield et al. 2001). Due to its negative regulation by H-NS and the fact it can complement *hns* mutant phenotypes it has been proposed that StpA may serve as a molecular “backup” for H-NS (Sonnenfield et al. 2001). Since H-NS expression is more or less constitutive it remains unclear what the necessity for such a backup system would be (Free and Dorman 1995).

There is some evidence that points to a function for StpA outside of being an H-NS backup. Although phenotypes for *stpA* mutants can be observed in the absence of H-NS (Deighan et al. 2000; Johansson and Uhlin 1999) it is apparent that *stpA hns* double mutant phenotypes are not always a simple exacerbation of *hns* phenotypes (Bertin et al. 2001). Further complicating matters, the effect of StpA on certain *hns* mutant phenotypes, like an effect on *bgl* expression, greatly depend on the particular mutant allele of *hns* being studied and this relates to the ability of StpA to interact with the mutant H-NS protein (Dersch et al. 1994; Free et al. 1998, 2001; Johansson et al. 2001; Ohta et al. 1999; Wolf et al. 2006). As outlined in Table 13.1, several H-NS mutants have been isolated that upregulate expression at the *proU* locus but only have a minimal effect on the *bgl* locus (Dersch et al. 1994; Ueguchi et al. 1996, 1997). Similar allele dependent effects have been observed for cold-sensitivity, mucoidy, and the regulation of various promoters (Free et al. 2001). The specific class of *hns* mutants that retain partial function lack DNA binding ability but retain a functional N-terminal domain capable of interacting with StpA. Studies have revealed that the truncated H-NS N-terminal domain interacts with StpA to silence the *bgl* operon with the StpA protein acting as the module that supplies the missing DNA-binding function (Free et al. 2001; Wolf et al. 2006). However, when the N-terminal domain of H-NS is expressed at very high-levels some repressive effects can be observed even in the absence of StpA by an unknown mechanism perhaps involving σ^S and Crl (Free et al. 2001; Ohta et al. 1999; Schnetz 2002). The role of the H-NS N-terminal domain in coordinating *bgl* silencing with StpA remains unclear but experiments suggest that the effect is not entirely by enhancing the stability of StpA (Wolf et al. 2006). Instead it appears that the N-terminal H-NS domain may alter StpA-nucleoprotein complex in a manner that facilitates silencing at the *bgl* operon. This result is somewhat surprising given that overexpression of H-NS N-terminal domains usually leads to anti-silencing at H-NS regulated promoters and also

Table 13.1 Mutations in *hns* and some of their reported phenotypes

Point mutations	
M1I	Increases in <i>fimB</i> and <i>proU</i> expression, <i>fimA</i> promoter inversion mutant (H-NS influence on promoter inversion is unknown), has nonmotile bacteria phenotype (Donato and Kawula 1999); express <i>bgl</i> and <i>proU</i> operons (Ueguchi et al. 1996)
R12C, R12H, R12A	Expresses <i>bgl</i> and <i>proU</i> operons, can bind DNA and <i>proU</i> similar to wild-type H-NS, oligomerization same as wild-type, cannot repress <i>proU</i> transcription in vitro (Ueguchi et al. 1996). Mutation renders the N-terminal H-NS fragment unable to bind Hha (Garcia et al. 2006). When combined with R15A mutation the protein displays lower DNA binding affinity and fails to distinguish a model curved sequence in vitro from a control fragment (Bloch et al. 2003)
R15C, R15H	Classified as a “class I” mutation by Ueguchi and colleagues because the mutant expresses both <i>bgl</i> and <i>proU</i> (Ueguchi et al. 1996). Significant loss in the ability of the N-terminal domain to dimerize (Garcia et al. 2006). When combined with the R12A mutation the protein displays lower DNA binding affinity and fails to distinguish a model curved sequence in vitro from a control fragment (Bloch et al. 2003)
A18E	Mutant induces <i>fimB</i> and <i>proU</i> expression, <i>fimA</i> promoter inversion mutant, hypermotile bacteria phenotype, only one H-NS isoform present on 2D electrophoresis (wild-type has 2 isoforms) (Donato and Kawula 1999)
C21F	Has no effect on stability of H-NS (Johansson et al. 2001); note: a similar mutation (C21S) in H-NS from <i>Vibrio</i> does not affect circular dichroism spec or melting characteristics or DNA-binding properties (Nye and Taylor 2003). Indicates that disulfide bond formation does not contribute to oligomerization
L26P	Mutant induces <i>fimB</i> and <i>proU</i> expression, <i>fimA</i> promoter inversion mutant, only one H-NS isoform present on 2D electrophoresis (wt 2 isoforms) (Donato and Kawula 1999); not functional and acts as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996)
L30D, L30P	These mutants when introduced into the N-terminal domain (residues 1-91) fail to heterooligomerize with wild-type H-NS and therefore cannot act in a dominant negative fashion on H-NS function. These mutations when introduced in full length H-NS cause derepression of <i>proU</i> and <i>bgl</i> (Ueguchi et al. 1997); L30P specifically disrupts oligomerization (Nye and Taylor 2003)
L30A, L30K	These mutants when introduced into the N-terminal domain (residues 1-91) still allow it to heterooligomerize with wild-type H-NS and act in a dominant negative fashion on H-NS function. These mutations in full length H-NS can repress represses <i>proU</i> and <i>bgl</i> in an <i>hns</i> background and therefore appear to have no effect on H-NS function (Ueguchi et al. 1997)
Δ A46 (<i>hns-107</i>)	Increased <i>proU</i> expression, no change in doubling time, no change in osmo-sensitivity, non-motile, slight increase in negative supercoiling (Hinton et al. 1992)
E53G/T55P	This mutant is poorly functional and acts as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996); footprints the <i>proU</i> promoter similar to wild-type H-NS at low concentrations but not at higher concentrations (Badaut et al. 2002)

(continued)

Table 13.1 (continued)

Point mutations	
R54C	Classified as a “class I” mutation by Ueguchi and colleagues because the mutant expresses both <i>bgl</i> and <i>proU</i> (Ueguchi et al. 1996)
L65P	Increases <i>fimB</i> and <i>proU</i> expression, <i>fimA</i> promoter inversion mutant (Donato and Kawula 1999)
R90H, R90C	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
A91T	Derepression of <i>proU</i> but not <i>bgl</i> , mutation was grouped as class II (Ueguchi et al. 1996)
R93C	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
R93H	Mutants induce <i>fimB</i> and <i>proU</i> expression (Donato and Kawula 1999)
P94L, P94S	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
A95T	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
Y97C, Y97H, Y97S	These mutants are poorly functional and act as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996); Y97C retains some specificity toward the <i>proU</i> promoter fragment although much reduced as compared to wild-type H-NS (Badaut et al. 2002)
T108I	Induces <i>fimB</i> and <i>proU</i> expression, results in hypermotile bacteria, defective in binding <i>fimB</i> DNA (Donato and Kawula 1999)
T110A	Mutant is poorly functional and acts as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996)
T110I	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
G111S, G111D	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996); G111S induces <i>fimB</i> and <i>proU</i> expression (Donato and Kawula 1999)
ΔG113-P116	Impaired ability to compact chromosome, not lethal when overexpressed (Spurio et al. 1997)
P116S, P116A, ΔP116	Mutant is poorly functional and acts as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996); unable to discriminate between curved and non-curved DNA and unable to bend DNA, also not toxic to cells when overexpressed (Spurio et al. 1997); P116S retains some specificity toward <i>proU</i> promoter over other DNA fragments although much reduced compared to wild-type H-NS (Badaut et al. 2002); P116S fails to promote intermolecular transposition of Tn10 (Ward et al. 2007)
G113D	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> – also shown to be less efficient at binding DNA and cannot repress <i>proU</i> transcription in vitro, but displays an enhanced ability to oligomerize (Ueguchi et al. 1996)
R114C	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
R114H	Results in lethality when overexpressed indicating that this mutant is still functional (Spurio et al. 1997)
T115I	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)

(continued)

Table 13.1 (continued)

Point mutations	
T115S	Results in lethality when overexpressed indicating that this mutant is still functional (Spurio et al. 1997)
I119T	Mutant is poorly functional and acts as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996); displays poor DNA binding and complete loss of specificity for <i>bla</i> and <i>proU</i> promoters compared to other DNA fragments (Badaut et al. 2002)
F133S	Mutant is poorly functional and acts as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996)
Truncations	
1–37 (<i>hns-206</i>)	Ampicillin-resistance cassette after codon 37 leads to a null allele that cannot be detected by immunoblot and fails to repress an ectopic <i>hns-lacZ</i> promoter fusion (Dersch et al. 1993); leads to cold sensitivity (Dersch et al. 1994); fails to silence <i>appY</i> and <i>bgl</i> promoters (Atlung et al. 1996; Dole et al. 2004a); causes reduced chromosomal ploidy (Atlung and Hansen 2002); leads to reduced <i>sbmC</i> expression (Oh et al. 2001); leads to increased levels of σ^S (Barth et al. 1995); fails to repress F-plasmid transfer genes (Will et al. 2004)
1–63 (mutant “N63”)	Mutant has a dominant negative effect in vivo: when expressed with wt H-NS, heterooligomerizes with wt and derepresses <i>proU</i> . Can form heterooligomers in vitro as detected by chemical cross-linking with DMS. (Ueguchi et al. 1997)
1–63	A frameshift after residue 63 resulting in an additional small sequence (NADR) followed by a stop codon, this mutation results in derepression of <i>fimB</i> and <i>proU</i> expression, <i>fimA</i> promoter inversion mutant (Donato and Kawula 1999)
1–64 ($\Delta 64C$)	A frameshift from loss of AT base-pair resulting in residues 1–64 followed by Cys and a stop codon. Not functional and acts as a dominant negative for <i>proU</i> expression (Williams et al. 1996)
1–82 (mutant “N82”)	Mutant induces <i>fimB</i> and <i>proU</i> expression (Donato and Kawula 1999)
1–87 (<i>hns-101</i>)	An IS10 insertion in K87 codon. Results in derepression of <i>proU</i> , increased doubling time, osmo-sensitivity (no growth on 0.3M NaCl), decreased motility, increased negative supercoiling (Hinton et al. 1992)
1–89 (<i>hns-103</i>)	IS10 insertion following K89 results in increased <i>proU</i> expression (Hinton et al. 1992)
1–91 (<i>hns60</i> , mutant “N91” or “Q92am”)	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> – also shown to be less efficient at binding DNA and cannot repress <i>proU</i> transcription in vitro, but displays an enhanced ability to oligomerize (Ueguchi et al. 1996, 1997); retains partial ability to repress <i>bgl</i> in a manner that requires σ^S and/or Crl (Free et al. 2001; Ohta et al. 1999; Schnetz 2002)
1–93 (mutant “N93”)	This mutation causes derepression of <i>fimB</i> and <i>proU</i> expression, <i>fimA</i> promoter inversion mutant, displays a two-fold decrease in motility (Donato and Kawula 1999)

(continued)

Table 13.1 (continued)

Point mutations	
1–93 (<i>hns-205</i>)	A Tn10 insertion in codon 93. Partially defective for cold adaptation but not to the extent of the <i>hns-206</i> allele (see above) (Dersch et al. 1994); fails to silence <i>appY</i> and <i>dsrA</i> expression (Atlung et al. 1996; Sledjeski and Gottesman 1995); causes reduced chromosomal ploidy (Atlung and Hansen 2002); reduced ability to stimulate σ^S expression in response to osmolarity (Barth et al. 1995); retains partial silencing of <i>bgl</i> in a StpA-dependent manner
1–108 (mutant “W109am”)	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
1–111 (mutant “Q112oc”)	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
1–121 (<i>hnsΔ121</i>)	A frameshift resulting in new sequence after K121 (KKQWMRKVPSTIS) ending with a TGA stop codon. This altered H-NS is poorly functional and acts as a dominant negative for expression of <i>proU</i> and semisynthetic 5A6A promoters (Williams et al. 1996)
1–123	A two base pair deletion resulting in a frameshift and an early stop codon. Induces <i>fimB</i> and <i>proU</i> expression; <i>fimA</i> promoter inversion mutant (Donato and Kawula 1999)
1–124	A one base pair insertion resulting in a frameshift and an early stop codon. Induces <i>fimB</i> and <i>proU</i> expression, three-fold decrease in motility (Donato and Kawula 1999)
1–125 (mutant “Q126oc”)	Classified as a “class II” mutation by Ueguchi and colleagues because the mutant expresses <i>proU</i> but not <i>bgl</i> (Ueguchi et al. 1996)
1–126 (<i>hns-102</i>)	An IS10 insertion in codon for Q126—results in increased <i>proU</i> expression, increased doubling time, osmo-sensitive (no growth on 0.3M NaCl), displays a motility defect, increased negative supercoiling (Hinton et al. 1992)
Δ2–20 (mutant “C118”)	Has a dominant negative effect in vivo when expressed with wild-type H-NS. Can heterooligomerize with H-NS and derepress <i>proU</i> . Heterooligomers detected with gel filtration (not detected with chemical cross-linking via DMS) (Ueguchi et al. 1997)
Δ2–69 (mutant “C69”)	Does not display dominant negative effect when expressed in background with wild-type H-NS (Ueguchi et al. 1997)

StpA-mediated silencing at some promoters (Williams et al. 1996). The effect that H-NS interaction has on StpA gene regulation is specific for a subset of promoters and does not seem to be required for silencing of the *stpA* promoter, for example (Wolf et al. 2006).

Experiments using mutants that generate truncated H-NS molecules are highly artificial and give very little information as to the possible role of StpA in wild-type cells. StpA expression is enhanced during the SOS response and is also affected by growth phase, supercoiling, temperature, osmolarity, and Lrp, suggesting it may play a role in regulating adaptation or stress resistance but the pathway by which it would do so is unclear (Deighan et al. 2003; Sonnenfield et al. 2001). StpA was originally isolated as a protein that, when overexpressed, could suppress a splicing-defective phage T4 *td* intron mutant in vivo. In vitro analysis revealed that it can act

as an RNA chaperone that enhances *td* intron splicing by preventing RNA misfolding and promoting formation of the active self-splicing intron through interactions via its C-terminal domain (Cusick and Belfort 1998; Zhang and Belfort 1992; Zhang et al. 1995, 1996). Both H-NS and StpA have been shown to bind RNA in vitro although StpA has a markedly higher affinity for RNA and is a much better chaperone for the *td* intron than H-NS (Brescia et al. 2004; Cusick and Belfort 1998; Zhang et al. 1996).

Although there is little evidence that StpA facilitates T4 *td* intron splicing during phage infection of wild-type cells, some evidence points to a potential role for StpA as an effector of RNA stability in another context. H-NS and StpA have each been implicated as playing a role in modulating RNA stability to control the expression of genes involved both in adaptation and stress resistance (Brescia et al. 2004; Deighan et al. 2000; Graeme-Cook et al. 1989; Suzuki et al. 1996). H-NS has been shown to also have RNA chaperone activity and may be directly involved in altering the stability of the *rpoS* message (Brescia et al. 2004). StpA can destabilize the small *micF* RNA involved in preventing the expression of the *ompF* transcript, although this effect was not observed in the presence of wild-type H-NS (Deighan et al. 2000). Transcription of *MicF* RNA expression is silenced by H-NS but activated by several regulators in response to a variety of stresses including antimicrobial peptides, oxidative stress, weak acids and changes in osmolarity (Delihias and Forst 2001). A role for StpA as a chaperone of small RNAs may explain the finding that silencing of the arginine decarboxylase gene in *hns* mutants can be restored by overexpression of either StpA or the small RNA binding protein Hfq (Shi and Bennett 1994). Studies attempting to delineate StpA function may be confounded by the fact they were generally carried out in *hns* mutants or using purified StpA protein, neither of which may be biologically relevant given the data that suggests StpA exists as an H-NS/StpA heterodimer in vivo.

13.12.2 Sfh – A Protein That Aids Transmission of Mobile Genetic Elements

As alluded to in earlier sections, several conjugative plasmids have been identified that encode H-NS and/or Hha homologues including the IncM plasmid R446, the IncN plasmid R46/pKM101, and members of the IncHI1 family of R27-like plasmids (Tietze and Tschäpe 1994). The most well studied plasmid-encoded H-NS homologues are the Sfh (*Shigella flexneri* H-NS-like protein) of plasmid pSF-R27 from *Shigella flexneri* 2a strain 2457T and the nearly identical ORF162 from the R27 plasmid of *Salmonella enterica* serovar Typhi (*S. typhi*) (Beloin et al. 2003; Forns et al. 2005). In terms of primary sequence these proteins actually bear a closer resemblance to StpA than to H-NS (Beloin et al. 2003). Like StpA, Sfh is capable of restoring wild-type expression of *proU* and *fliC* and complementing Bgl, mucoidy, and porin protein expression phenotypes when expressed in *E. coli hns* mutants. A similar backup function has been observed for ORF162 encoded on the

R27 plasmid for *S. typhi* (Forns et al. 2005). The functional equivalence of Sfh has further been demonstrated by the fact that the Sfh N-terminal domain can form heterodimers with StpA or H-NS in yeast two-hybrid assays (Deighan et al. 2003). StpA can both transcriptionally regulate and be regulated by other H-NS family members although in wild-type cells it appears to have little effect on expression of endogenous H-NS. Regulation of Sfh is highly unusual in that transcript levels are highest during logarithmic growth but protein does not accumulate at appreciable levels until early stationary phase (Deighan et al. 2003). The underlying mechanism behind the reciprocal regulation of Sfh expression is unclear but it seems to involve a blockade of Sfh translation during log phase growth that is relieved during *Shigella* entry into stationary phase (Doyle and Dorman 2006).

Given that plasmid-encoded H-NS-like molecules are frequently associated with conjugative plasmids, and the fact that H-NS can regulate conjugation in many systems (see Section 13.10.1), it is possible that one role of Sfh is to regulate the conjugative apparatus of pSF-R27. The plasmid-encoded homologs may enable proper regulation if the plasmid transfers to a cell that lacks H-NS or they may respond to certain environmental conditions that the endogenous H-NS molecules do not. This idea is supported by the finding that deletion of either ORF162, the R27 encoded H-NS homologue of *S. typhi*, or ORF181 (an R27 plasmid encoded Hha homologue) greatly increases conjugation frequency at elevated temperatures (Forns et al. 2005). The effects of deleting ORF162 and ORF181 are insignificant at 25°C, where conjugation frequencies are already elevated. However, the finding that plasmid encoded H-NS homologues have a function that cannot necessarily be replaced by H-NS suggests that there may be subtle differences between H-NS and the plasmid encoded H-NS-like molecules.

A different hypothesis has recently been advanced to explain why mobile genetic elements encode H-NS-like molecules. Doyle et al. observed that transfer of pSF-R27 lacking *sfh* to *S. typhimurium* strain 14028s had a profound effect on the *Salmonella* transcriptome, altering the expression of over 400 genes while transcription of fewer than 100 *Salmonella* genes were affected by transfer of the wild-type plasmid (Doyle et al. 2007). The affected genes covered a broad range of physiological categories and a surprisingly large number were involved in adaptation and stress resistance. Accordingly, the *Salmonella* strain harboring the Δsfh plasmid demonstrated enhanced UV resistance and an increased ability to survive in cultured J774 macrophages. *Salmonella* harboring the Δsfh plasmid displayed a significant fitness defect during growth in liquid media compared to wild-type strains. A small number of H-NS silenced genes were misregulated in the presence of the Δsfh plasmid and bacterial motility was severely reduced. Importantly, all of these phenotypes could be relieved by providing *sfh* or *hns* or by augmenting expression of H-NS. In another experiment it was found that a very high copy number plasmid harboring the *ssrA* promoter also caused a fitness defect compared to wild type and this fitness defect could be relieved by augmented expression of either Sfh or H-NS as well.

These findings led to the hypothesis that the fitness and motility defects observed in the *Salmonella* strain harboring the mutant plasmid were due to a titration of the

endogenous cellular pools of H-NS (Doyle et al. 2007). Further experimentation is necessary to address several potential issues with the titration model. First, the majority of the *Salmonella* genes whose expression was altered upon introduction of the Δsfh plasmid were not identified as H-NS-regulated in other studies and many known H-NS regulated genes were unaffected. Many of the genes found to be altered were involved in stress resistance suggesting that Δsfh plasmid may have more to do with an adverse alteration of cell physiology than on titration of H-NS per se (Doyle et al. 2007). A possible role for Sfh in preventing stress during conjugal transfer is consistent with findings that conjugative plasmids have been shown to induce transient stress responses including the SOS response, particularly in the recipient cell (Althorpe et al. 1999; Bailone et al. 1988; Golub et al. 1988). Second, expression of H-NS is autoregulated and it is unlikely, though possible, that H-NS cannot increase its production to compensate for the relatively small amount of AT-rich sequence carried by the low-copy pSF-R27 plasmid (Dersch et al. 1993; Falconi et al. 1991; Free and Dorman 1995; Ueguchi et al. 1993). Third, many naturally occurring plasmids that lack H-NS homologues have no apparent fitness defect on the cell, although no detailed analysis has been performed to correlate the amount of AT-rich sequence present on a plasmid to the likelihood that it encodes an H-NS-like protein. Finally, although the Δsfh plasmid had strong effects on the transcriptome when transferred to *Salmonella*, the Δsfh mutation has little, if any effect on virulence gene expression in *Shigella*, suggesting that the observed fitness defect may be species/strain specific or due to the altered expression of another factor present on the plasmid (Beloïn et al. 2003; Deighan et al. 2003). A complete transcriptional profile has not been performed for the Δsfh strain of *Shigella*, however, and it is possible that changes in the expression of many genes were not detected. Regardless of the specific mechanism, it is clear the Sfh protein can buffer fitness defect due to the presence of pSF-R27 and therefore plasmid-encoded H-NS like molecules likely have profound consequences in enabling conjugative plasmids to propagate themselves among varied bacterial populations.

13.12.3 *Pseudomonas MvaT/MvaU*

MvaT and MvaU (MvaT2) are paralogous DNA-bridging proteins in *Pseudomonas* species that have both structural and functional similarity to H-NS despite sharing less than 20% identity with H-NS in primary sequence (Dame et al. 2005; Tendeng et al. 2003b). MvaT expression in *E. coli hns* mutants restores motility and growth on minimal media supplemented with serine and blocks the ability of *hns* mutants to utilize salicin as a sole carbon source (Tendeng et al. 2003b). MvaT has been shown to be a negative regulator of the chaperone usher pathway (*cup*) clusters that encode pili essential for the formation of biofilms (Vallet et al. 2004), genes involved in quorum sensing and virulence including *rpoS* and *lecA* (Diggle et al. 2002, 2003), as well as the *mexAB-oprN* operon encoding a drug pump involved in resistance to chloramphenicol, imipenim and norfloxacin (Westfall et al. 2006).

The diversity of *mvaT* phenotypes and its involvement in quorum sensing led to the speculation that MvaT was a global regulator of gene expression (Diggle et al. 2002). However, in contrast to what has been observed with H-NS, microarray analysis of *Pseudomonas aeruginosa mvaT* mutants found that the number of genes under control of MvaT is relatively small; approximately 100 genes very few of which demonstrate greater than 5-fold downregulation by MvaT (Vallet et al. 2004). Recent findings have determined that a primary reason that *Pseudomonas mvaT* mutants are not as phenotypically remarkable as *hns* mutants in the enteric bacteria is that MvaT and MvaU share very strong functional redundancy. While mutations in either *mvaT* or *mvaU* are tolerated relatively well by *P. aeruginosa*, they are lethal for the bacteria when combined (Castang et al. 2008). ChIP-on-chip experiments to determine the genome-wide binding sites for MvaT and MvaU have revealed nearly complete overlap in their respective binding profiles including a strong preference for AT-rich patches in the genome that is highly reminiscent of what has been determined for H-NS. These findings suggest that MvaT and MvaU play a role in the silencing of xenogeneic genes in *Pseudomonas sp.* although this has yet to be formally demonstrated.

13.12.4 *Bordetella BpH3*

BpH3, the first H-NS-like molecule discovered outside of the *Enterobacteriaceae*, was originally isolated from *Bordetella pertussis* as a 16 kDa DNA-binding protein through southwestern blotting analysis (Goyard 1996). BpH3 is approximately 30% identical to H-NS with the strongest homology occurring within the C-terminal domain (Goyard and Bertin 1997). Expression of BpH3 in an *E. coli* strain lacking H-NS restores motility and high-level expression of the *flhDC* and *bgl* operons (Goyard and Bertin 1997). Competitive gel retardation assays show that, like H-NS, BpH3 displays higher binding affinity for the *bla* and *flhDC* promoter fragments compared to other sequences contained within expression plasmids (Goyard and Bertin 1997). Homologues of BpH3 are present in *B. bronchiseptica* (BbH3) and *B. parapertussis* but the function of these proteins in their native context has not been explored.

13.12.5 *Mycobacterial Lsr2*

Lsr2 is a small and basic DNA-binding protein conserved among the *Mycobacteria* and *Actinomycetes*. Although it bears no sequence similarity to H-NS, recent evidence suggests that Lsr2 is a pleiotropic regulator that shares several properties in common with H-NS. *Mycobacterium smegmatis lsr2* mutants display a smooth colony morphology, enhanced sliding motility, increased phage resistance, and an altered ability to form biofilms (Arora et al. 2008; Chen et al. 2006; Colangeli et al. 2007;

Kocincova et al. 2008). These defects may be due, in part, to an alteration in surface lipids including a loss of mycolyl-diacylglycerols (Chen et al. 2006), an increase in glycopeptidolipids (Kocincova et al. 2008), or an increase in another unidentified polar lipid (Colangeli et al. 2007), although these findings have been somewhat controversial and may be strain specific (Arora et al. 2008). Lsr2 has further been implicated as a repressor of the *iniBAC* genes involved in mycobacterial drug resistance and, indeed, loss of Lsr2 leads to increased resistance to ethambutol (Colangeli et al. 2007) and kanamycin (Arora et al. 2008).

Lsr2 exists as a homodimer that can bridge DNA in a manner that suggests cooperativity and appears to share a basic domain structure similar to that of H-NS (Chen et al. 2008). In vitro binding studies indicate that supercoiled DNA is a preferred target over linear templates, but that curved DNA is not preferentially bound over non-curved sequences (Chen et al. 2008; Colangeli et al. 2007). Microarray analysis suggests that Lsr2 acts primarily as a negative regulator and that most targeted genes have distinctly AT-rich 5' untranslated sequences (Colangeli et al. 2007). Highlighting the fact that H-NS and Lsr2 may be functionally equivalent it has recently been shown that expression of H-NS can restore wild-type colony morphology in *M. smegmatis* *lsr2* mutants and that Lsr2 expression in *E. coli* *hns* mutants restores motility and represses hemolysis and expression of the *bgl* operon (Gordon et al. 2008). Chromatin immunoprecipitation analysis of Lsr2 binding in *E. coli* *hns* mutants reveals that AT-rich genes are specifically targeted and that overexpression of SlyA, which antagonizes H-NS at the *hlyE* promoter, is also able to counteract Lsr2-mediated repression at the *hlyE* locus.

The fact that Lsr2 bears considerable functional correspondence to H-NS without sharing any similarity in the primary sequence is exciting on many levels. It suggests that the universe of "H-NS like" molecules may be significantly larger than has been revealed through phylogenetic analysis based on sequence similarity. It remains to be determined if the common functionality of Lsr2 and H-NS is the result of convergent evolution and whether species other than the high-GC Gram-positive bacteria and the subset of α -, β - and γ -proteobacteria have DNA-bridging proteins that can target AT-rich sequences in a non-sequence specific manner.

13.13 Directions for the Future/Unanswered Questions

For all that we now understand about H-NS, how it binds to DNA, and the genes under its control it is striking how many aspects of the basic features of this molecule remain controversial or completely unknown. This is not due to a lack of effort on the part of many groups, but rather it reflects the fact that the data obtained from different methods have been at times completely contradictory. Although the recent availability of both high-throughput and single molecule studies have greatly improved our understanding of H-NS the molecule has stubbornly resisted all attempts to clarify many of its important features. I list a few of the remaining unresolved areas below:

Orientation of the dimerization domain: The H-NS dimerization domain has been reported to be arranged in either a parallel or anti-parallel (handshake) fashion depending on the particular study. This discrepancy is remarkable given that the basic structures of the H-NS monomers are similar in both reported structures. Is it possible that both interpretations are correct? One study has proposed that the functional unit is a tetramer with the four molecules arranged in a combination of anti-parallel and parallel orientations (Stella et al. 2005). This hybrid model seems unlikely since critical residues involved in dimerization in the parallel structure are also critical for formation of the anti-parallel structure. Furthermore the dimer surfaces in the two proposed arrangements are predicted to be negatively charged which also would appear to refute the idea that a tetramer of this nature is possible.

Mechanism of cooperativity: We still have no clear mechanistic understanding of the mechanism by which cooperativity is achieved. Most reports using traditional biochemical techniques suggested H-NS dimers interact with one another to make higher order complexes such as tetramers or longer filaments. Recent experiments with the Q-trap suggest that the functional unit is solely a dimer and that apparent cooperativity may be merely a function of the fact that the first binding event brings the strands closer together, thereby lowering the entropic barrier to subsequent binding events (Dame et al. 2006). This model is also supported by the structural study of the dimerization domain performed by Bloch et. al., which suggested that the higher-order H-NS complexes observed in solution are merely non-specific aggregates (Bloch et al. 2003). A combined use of Q-trap and previously defined mutants in the oligomerization domain may be informative in determining whether higher-order H-NS complexes actually form through protein-protein interaction when bound to DNA.

Mechanism of binding and silencing: It remains unclear what sequence elements are specifically recognized by H-NS. Does H-NS specifically target a limited set of consensus sequences such as the consensus recently derived from *proU* or does it recognize more general features including curvature or AT-richness? Some studies have suggested that H-NS has the ability to bind DNA in both a specific and non-specific manner, but is the “non-specific” mechanism truly relevant in vivo? There are also indirect observations that suggest bridging itself may not be the universal mechanism by which H-NS binds DNA and some evidence suggests there may be a second mode whereby H-NS binds without bridging adjacent strands. Would this non-bridging mode of binding differ in resulting regulation or ability to constrain supercoils?

The origins and functions of H-NS homologues: H-NS-like molecules have a bizarre distribution among Gram-negative bacterial species and they are often located on mobile genetic elements or conjugative plasmids. Does H-NS act as a silencer in most cases or, as in the case of *Ler*, do many of these molecules have the ability to bind specific sequences to carry out specific functions at a limited set of genes? Is it safe to assume that H-NS paralogues have functions other than acting as a “backup” for H-NS and, if so, what are the biological functions of StpA and H-NS-like molecules that are plasmid encoded? Does the high degree of sequence

diversity between H-NS homologues indicate there has been an evolutionary arms race between mobile genetic elements and phages against H-NS?

The role of Hha-like molecules: In the *Enterobacteriaceae* it is clear that the Hha-like molecules play an important role as a co-factor that allows H-NS to effectively silence expression from large numbers of genes. We currently lack any understanding of the underlying mechanism behind this phenomenon and do not know which loci are directly Hha-regulated and which effects are indirect. Chromatin immunoprecipitation studies should help in this regard but will likely leave unresolved the question of how and why Hha-like molecules facilitate silencing, why they share so much functional redundancy, and why their distribution is limited to the *Enterobacteriaceae*.

It is clear that the answers to many of these questions will not be uncovered without the development of new tools or through the application of significantly new approaches. Just the few examples explored in detail thus far, such as the *bgl* and *proU* operons, reveal that higher-order promoter geometry plays an important role in determining whether H-NS/Hha will be effective in silencing. Significant understanding of the silencing mechanism will not occur without the ability to visualize promoter geometry directly and under conditions that effectively mimic those found *in vivo*. Recent advances in single (or dual) molecule manipulation in conjunction with the use of the well-defined promoter and protein mutations that were uncovered with more traditional approaches will hopefully shed some light on these important biological issues.

Acknowledgements I would like to acknowledge the help of Sabrina Ali and Robin Imperial in assembling the table of published *hns* mutants and their phenotypes. I would further like to thank David Haniford, Blair Gordon, and Simon Dove for critical reading of portions of the manuscript.

References

- Abe H, Miyahara A, Oshima T, Tashiro K, Ogura Y, Kuhara S, Ogasawara N, Hayashi T, Tobe T (2008) Global regulation by horizontally transferred regulators establishes the pathogenicity of *Escherichia coli*. *DNA Res* 15:25–38
- Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181:6361–6370
- Althorpe NJ, Chille PM, Thomas AT, Brammar WJ, Wilkins BM (1999) Transient transcriptional activation of the Inc11 plasmid anti-restriction gene (*ardA*) and SOS inhibition gene (*psiB*) early in conjugating recipient bacteria. *Mol Microbiol* 31:133–142
- Altuvia S, Almiron M, Huisman G, Kolter R, Storz G (1994) The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol Microbiol* 13:265–272
- Amit R, Oppenheim AB, Stavans J (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophys J* 84:2467–2473
- Amit R, Oppenheim AB, Stavans J (2004) Single molecule elasticity measurements: a biophysical approach to bacterial nucleoid organization. *Biophys J* 87:1392–1393
- Arnqvist A, Olsen A, Normark S (1994) Sigma S-dependent growth-phase induction of the *csgBA* promoter in *Escherichia coli* can be achieved *in vivo* by sigma 70 in the absence of the nucleoid-associated protein H-NS. *Mol Microbiol* 13:1021–1032

- Arora K, Whiteford DC, Lau-Bonilla D, Davitt CM, Dahl JL (2008) Inactivation of *lsr2* results in a hypermotile phenotype in *Mycobacterium smegmatis*. *J Bacteriol* 190:4291–4300
- Atlung T, Hansen FG (2002) Effect of different concentrations of H-NS protein on chromosome replication and the cell cycle in *Escherichia coli*. *J Bacteriol* 184:1843–1850
- Atlung T, Ingmer H (1997) H-NS: a modulator of environmentally regulated gene expression. *Mol Microbiol* 24:7–17
- Atlung T, Sund S, Olesen K, Brondsted L (1996) The histone-like protein H-NS acts as a transcriptional repressor for expression of the anaerobic and growth phase activator AppY of *Escherichia coli*. *J Bacteriol* 178:3418–3425
- Badaut C, Williams R, Arluison V, Bouffartigues E, Robert B, Buc H, Rimsky S (2002) The degree of oligomerization of the H-NS nucleoid structuring protein is related to specific binding to DNA. *J Biol Chem* 277:41657–41666
- Bae SH, Liu D, Lim HM, Lee Y, Choi BS (2008) Structure of the nucleoid-associated protein *cnu* reveals common binding sites for H-NS in *cnu* and *hha*. *Biochemistry* 47:1993–2001
- Bailone A, Backman A, Sommer S, Celerier J, Bagdasarian MM, Bagdasarian M, Devoret R (1988) PsiB polypeptide prevents activation of RecA protein in *Escherichia coli*. *Mol Gen Genet* 214:389–395
- Bakaev VV (1981) Structure of chromosomal deoxyribonucleoproteins. XI. Organization of deoxyribonucleoprotein complex in bacterial cells. *Mol Biol (Mosk)* 15:1350–1363
- Balsalobre C, Juarez A, Madrid C, Mourino M, Prenafeta A, Munoz FJ (1996) Complementation of the *hha* mutation in *Escherichia coli* by the *ymoA* gene from *Yersinia enterocolitica*: dependence on the gene dosage. *Microbiology* 142:1841–1846
- Banos RC, Pons JJ, Madrid C, Juarez A (2008) A global modulatory role for the *Yersinia enterocolitica* H-NS protein. *Microbiology* 154:1281–1289
- Barba J, Bustamante VH, Flores-Valdez MA, Deng W, Finlay BB, Puente JL (2005) A positive regulatory loop controls expression of the locus of enterocyte effacement-encoded regulators *Ler* and *GrlA*. *J Bacteriol* 187:7918–7930
- Barth M, Marschall C, Muffler A, Fischer D, Hengge-Aronis R (1995) Role for the histone-like protein H-NS in growth phase-dependent and osmotic regulation of sigma S and many sigma S-dependent genes in *Escherichia coli*. *J Bacteriol* 177:3455–3464
- Becker G, Hengge-Aronis R (2001) What makes an *Escherichia coli* promoter sigma(S) dependent? Role of the –13/–14 nucleotide promoter positions and region 2.5 of sigma(S). *Mol Microbiol* 39:1153–1165
- Beloin C, Dorman CJ (2003) An extended role for the nucleoid structuring protein H-NS in the virulence gene regulatory cascade of *Shigella flexneri*. *Mol Microbiol* 47:825–838
- Beloin C, Deighan P, Doyle M, Dorman CJ (2003) *Shigella flexneri* 2a strain 2457T expresses three members of the H-NS-like protein family: characterization of the Sfh protein. *Mol Genet Genom* 270:66–77
- Berdichevsky T, Friedberg D, Nadler C, Rokney A, Oppenheim A, Rosenshine I (2005) *Ler* is a negative autoregulator of the LEE1 operon in enteropathogenic *Escherichia coli*. *J Bacteriol* 187:349–357
- Bertin P, Hommais F, Krin E, Soutourina O, Tendeng C, Derzelle S, Danchin A (2001) H-NS and H-NS-like proteins in Gram-negative bacteria and their multiple role in the regulation of bacterial metabolism. *Biochimie* 83:235–241
- Blanco A, Juarez A, Pastor FI (1991) Overproduction of a *Clostridium cellulolyticum* endoglucanase by mutant strains of *Escherichia coli*. *FEMS Microbiol Lett* 65:221–226
- Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453–1474
- Bloch V, Yang Y, Margeat E, Chavanieu A, Auge MT, Robert B, Arold S, Rimsky S, Kochoyan M (2003) The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat Struct Biol* 10:212–218
- Bouffartigues E, Buckle M, Badaut C, Travers A, Rimsky S (2007) H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat Struct Mol Biol* 14:441–448

- Bracco L, Kotlarz D, Kolb A, Diekmann S, Buc H (1989) Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. EMBO J 8:4289–4296
- Brescia CC, Kaw MK, Sledjeski DD (2004) The DNA binding protein H-NS binds to and alters the stability of RNA *in vitro* and *in vivo*. J Mol Biol 339:505–514
- Browning DF, Cole JA, Busby SJ (2000) Suppression of FNR-dependent transcription activation at the *Escherichia coli nir* promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. Mol Microbiol 37:1258–1269
- Bryan A, Roesch P, Davis L, Moritz R, Pellett S, Welch RA (2006) Regulation of type 1 fimbriae by unlinked FimB- and FimE-like recombinases in uropathogenic *Escherichia coli* strain CFT073. Infect Immun 74:1072–1083
- Bustamante VH, Santana FJ, Calva E, Puente JL (2001) Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression. Mol Microbiol 39:664–678
- Caramel A, Schnetz K (1998) Lac and lambda repressors relieve silencing of the *Escherichia coli bgl* promoter. Activation by alteration of a repressing nucleoprotein complex. J Mol Biol 284:875–883
- Caramel A, Schnetz K (2000) Antagonistic control of the *Escherichia coli bgl* promoter by FIS and CAP in vitro. Mol Microbiol 36:85–92
- Carmona M, Balsalobre C, Munoa F, Mourino M, Jubete Y, De la Cruz F, Juarez A (1993) *Escherichia coli hha* mutants, DNA supercoiling and expression of the haemolysin genes from the recombinant plasmid pANN202–312. Mol Microbiol 9:1011–1018
- Caron E, Crepin VF, Simpson N, Knutton S, Garmendia J, Frankel G (2006) Subversion of actin dynamics by EPEC and EHEC. Curr Opin Microbiol 9:40–45
- Castang S, McManus HR, Turner KH, Dove SL (2008) H-NS family members function coordinately in an opportunistic pathogen. Proc Natl Acad Sci USA 105:18947–18952
- Cerdan R, Bloch V, Yang Y, Bertin P, Dumas C, Rimsky S, Kochoyan M, Arold ST (2003) Crystal structure of the N-terminal dimerisation domain of VicH, the H-NS-like protein of *Vibrio cholerae*. J Mol Biol 334:179–185
- Ceschini S, Lupidi G, Coletta M, Pon CL, Fioretti E, Angeletti M (2000) Multimeric self-assembly equilibria involving the histone-like protein H-NS. A thermodynamic study. J Biol Chem 275:729–734
- Chen CC, Wu HY (2005) LeuO protein delimits the transcriptionally active and repressive domains on the bacterial chromosome. J Biol Chem 280:15111–15121
- Chen CC, Chou MY, Huang CH, Majumder A, Wu HY (2005) A cis-spreading nucleoprotein filament is responsible for the gene silencing activity found in the promoter relay mechanism. J Biol Chem 280:5101–5112
- Chen JM, German GJ, Alexander DC, Ren H, Tan T, Liu J (2006) Roles of Lsr2 in colony morphology and biofilm formation of *Mycobacterium smegmatis*. J Bacteriol 188:633–641
- Chen JM, Ren H, Shaw JE, Wang YJ, Li M, Leung AS, Tran V, Berbenetz NM, Kocincova D, Yip CM, Reytrat JM, Liu J (2008) Lsr2 of *Mycobacterium tuberculosis* is a DNA-bridging protein. Nucleic Acids Res 36:2123–2135
- Colangeli R, Helb D, Vilcheze C, Hazbon MH, Lee CG, Safi H, Sayers B, Sardone I, Jones MB, Fleischmann RD, Peterson SN, Jacobs WR, Jr., Alland D (2007) Transcriptional regulation of multi-drug tolerance and antibiotic-induced responses by the histone-like protein Lsr2 in *M. tuberculosis*. PLoS Pathog 3:e87
- Colonna B, Casalino M, Fradiani PA, Zagaglia C, Naitza S, Leoni L, Prosseda G, Coppo A, Ghelardini P, Nicoletti M (1995) H-NS regulation of virulence gene expression in enteroinvasive *Escherichia coli* harboring the virulence plasmid integrated into the host chromosome. J Bacteriol 177:4703–4712
- Connell I, Agace W, Klemm P, Schembri M, Marild S, Svanborg C (1996) Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. Proc Natl Acad Sci USA 93:9827–9832
- Coombes BK, Wickham ME, Lowden MJ, Brown NF, Finlay BB (2005) Negative regulation of *Salmonella* pathogenicity island 2 is required for contextual control of virulence during typhoid. Proc Natl Acad Sci USA 102:17460–17465

- Corbett D, Bennett HJ, Askar H, Green J, Roberts IS (2007) SlyA and H-NS regulate transcription of the *Escherichia coli* K5 capsule gene cluster, and expression of *slyA* in *Escherichia coli* is temperature-dependent, positively autoregulated, and independent of H-NS. *J Biol Chem* 282:33326–33335
- Cordeiro TN, Garci AJ, Pons JI, Aznar S, Juarez A, Pons M (2008) A single residue mutation in Hha preserving structure and binding to H-NS results in loss of H-NS mediated gene repression properties. *FEBS Lett* 582:3139–3144
- Cornelis GR, Sluiter C, Delor I, Geib D, Kaniga K, Lambert de Rouvroit C, Sory MP, Vanooteghem JC, Michiels T (1991) *ymoA*, a *Yersinia enterocolitica* chromosomal gene modulating the expression of virulence functions. *Mol Microbiol* 5:1023–1034
- Cukier-Kahn R, Jacquet M, Gros F (1972) Two heat-resistant, low molecular weight proteins from *Escherichia coli* that stimulate DNA-directed RNA synthesis. *Proc Natl Acad Sci USA* 69:3643–3647
- Cusick ME, Belfort M (1998) Domain structure and RNA annealing activity of the *Escherichia coli* regulatory protein StpA. *Mol Microbiol* 28:847–857
- Dame RT (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* 56:858–870
- Dame RT (2008) Single-molecule micromanipulation studies of DNA and architectural proteins. *Biochem Soc Trans* 36:732–737
- Dame RT, Goosen N (2002) HU: promoting or counteracting DNA compaction? *FEBS Lett* 529:151–156
- Dame RT, Wuite GJ (2003) On the role of H-NS in the organization of bacterial chromatin: from bulk to single molecules and back. *Biophys J* 85:4146–4148
- Dame RT, Wyman C, Goosen N (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* 28:3504–3510
- Dame RT, Wyman C, Goosen N (2001) Structural basis for preferential binding of H-NS to curved DNA. *Biochimie* 83:231–234
- Dame RT, Wyman C, Wurm R, Wagner R, Goosen N (2002) Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the *rrnB* P1. *J Biol Chem* 277:2146–2150
- Dame RT, Luijsterburg MS, Krin E, Bertin PN, Wagner R, Wuite GJ (2005) DNA bridging: a property shared among H-NS-like proteins. *J Bacteriol* 187:1845–1848
- Dame RT, Noom MC, Wuite GJ (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* 444:387–390
- Dattananda CS, Rajkumari K, Gowrishankar J (1991) Multiple mechanisms contribute to osmotic inducibility of *proU* operon expression in *Escherichia coli*: demonstration of two osmoresponse promoters and of a negative regulatory element within the first structural gene. *J Bacteriol* 173:7481–7490
- Daubin V, Lerat E, Perriere G (2003a) The source of laterally transferred genes in bacterial genomes. *Genome Biol* 4:R57
- Daubin V, Moran NA, Ochman H (2003b) Phylogenetics and the cohesion of bacterial genomes. *Science* 301:829–832
- De Biase D, Tramonti A, Bossa F, Visca P (1999) The response to stationary-phase stress conditions in *Escherichia coli*: role and regulation of the glutamic acid decarboxylase system. *Mol Microbiol* 32:1198–1211
- de la Cruz F, Carmona M, Juarez A (1992) The Hha protein from *Escherichia coli* is highly homologous to the YmoA protein from *Yersinia enterocolitica*. *Mol Microbiol* 6:3451–3452
- Defez R, De Felice M (1981) Cryptic operon for beta-glucoside metabolism in *Escherichia coli* K12: genetic evidence for a regulatory protein. *Genetics* 97:11–25
- Deighan P, Free A, Dorman CJ (2000) A role for the *Escherichia coli* H-NS-like protein StpA in OmpF porin expression through modulation of *micF* RNA stability. *Mol Microbiol* 38:126–139
- Deighan P, Beloin C, Dorman CJ (2003) Three-way interactions among the Sfh, StpA and H-NS nucleoid-structuring proteins of *Shigella flexneri* 2a strain 2457T. *Mol Microbiol* 48:1401–1416

- Delilhas N, Forst S (2001) MicF: an antisense RNA gene involved in response of *Escherichia coli* to global stress factors. *J Mol Biol* 313:1–12
- Deng S, Stein RA, Higgins NP (2005) Organization of supercoil domains and their reorganization by transcription. *Mol Microbiol* 57:1511–1521
- Dersch P, Schmidt K, Bremer E (1993) Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol Microbiol* 8:875–889
- Dersch P, Kneip S, Bremer E (1994) The nucleoid-associated DNA-binding protein H-NS is required for the efficient adaptation of *Escherichia coli* K-12 to a cold environment. *Mol Gen Genet* 245:255–259
- Diderichsen B (1980) *cur-1*, a mutation affecting a phenotype of *sup+* strains of *Escherichia coli*. *Mol Gen Genet* 180:425–428
- Diggle SP, Winzer K, Lazdunski A, Williams P, Camara M (2002) Advancing the quorum in *Pseudomonas aeruginosa*: MvaT and the regulation of N-acylhomoserine lactone production and virulence gene expression. *J Bacteriol* 184:2576–2586
- Diggle SP, Winzer K, Chhabra SR, Worrall KE, Camara M, Williams P (2003) The *Pseudomonas aeruginosa* quinolone signal molecule overcomes the cell density-dependency of the quorum sensing hierarchy, regulates *rhl*-dependent genes at the onset of stationary phase and can be produced in the absence of LasR. *Mol Microbiol* 50:29–43
- Dole S, Klingens Y, Nagarajavel V, Schnetz K (2004a) The protease Lon and the RNA-binding protein Hfq reduce silencing of the *Escherichia coli* *bgl* operon by H-NS. *J Bacteriol* 186:2708–2716
- Dole S, Nagarajavel V, Schnetz K (2004b) The histone-like nucleoid structuring protein H-NS represses the *Escherichia coli* *bgl* operon downstream of the promoter. *Mol Microbiol* 52:589–600
- Donato GM, Kawula TH (1999) Phenotypic analysis of random *hns* mutations differentiate DNA-binding activity from properties of *fimA* promoter inversion modulation and bacterial motility. *J Bacteriol* 181:941–948
- Donato GM, Lelivelt MJ, Kawula TH (1997) Promoter-specific repression of *fimB* expression by the *Escherichia coli* nucleoid-associated protein H-NS. *J Bacteriol* 179:6618–6625
- Dorman CJ (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* 2:391–400
- Dorman CJ (2006) DNA supercoiling and bacterial gene expression. *Sci Prog* 89:151–166
- Dorman CJ (2007) H-NS, the genome sentinel. *Nat Rev Microbiol* 5:157–161
- Doyle M, Dorman CJ (2006) Reciprocal transcriptional and posttranscriptional growth-phase-dependent expression of *sfh*, a gene that encodes a paralogue of the nucleoid-associated protein H-NS. *J Bacteriol* 188:7581–7591
- Doyle M, Fookes M, Ivens A, Mangan MW, Wain J, Dorman CJ (2007) An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science* 315:251–252
- Duong N, Osborne S, Bustamante VH, Tomljenovic AM, Puente JL, Coombes BK (2007) Thermosensing coordinates a cis-regulatory module for transcriptional activation of the intracellular virulence system in *Salmonella enterica* serovar Typhimurium. *J Biol Chem* 282:34077–34084
- Elliott SJ, Sperandio V, Giron JA, Shin S, Mellies JL, Wainwright L, Hutcheson SW, McDaniel TK, Kaper JB (2000) The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect Immun* 68:6115–6126
- Ellison DW, Miller VL (2006) H-NS Represses *inv* transcription in *Yersinia enterocolitica* through competition with RovA and interaction with YmoA. *J Bacteriol* 188:5101–5112
- Ellison DW, Young B, Nelson K, Miller VL (2003) YmoA negatively regulates expression of invasins from *Yersinia enterocolitica*. *J Bacteriol* 185:7153–7159
- Esposito D, Petrovic A, Harris R, Ono S, Eccleston JF, Mbabaali A, Haq I, Higgins CF, Hinton JC, Driscoll PC, Ladbury JE (2002) H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *J Mol Biol* 324:841–850

- Falconi M, Gualtieri MT, La Teana A, Losso MA, Pon CL (1988) Proteins from the prokaryotic nucleoid: primary and quaternary structure of the 15-kD *Escherichia coli* DNA binding protein H-NS. *Mol Microbiol* 2:323–329
- Falconi M, McGovern V, Gualerzi C, Hillyard D, Higgins NP (1991) Mutations altering chromosomal protein H-NS induce mini-Mu transposition. *New Biol* 3:615–625
- Falconi M, Brandi A, La Teana A, Gualerzi CO, Pon CL (1996) Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of *hns* expression. *Mol Microbiol* 19:965–975
- Falconi M, Colonna B, Prosseda G, Micheli G, Gualerzi CO (1998) Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. *EMBO J* 17:7033–7043
- Falconi M, Prosseda G, Giangrossi M, Beghetto E, Colonna B (2001) Involvement of FIS in the H-NS-mediated regulation of *virF* gene of *Shigella* and enteroinvasive *Escherichia coli*. *Mol Microbiol* 42:439–452
- Finkel SE, Johnson RC (1992) The Fis protein: it's not just for DNA inversion anymore. *Mol Microbiol* 6:3257–3265
- Fletcher SA, Csonka LN (1995) Fine-structure deletion analysis of the transcriptional silencer of the *proU* operon of *Salmonella typhimurium*. *J Bacteriol* 177:4508–4513
- Forns N, Banos RC, Balsalobre C, Juarez A, Madrid C (2005) Temperature-dependent conjugative transfer of R27: role of chromosome- and plasmid-encoded Hha and H-NS proteins. *J Bacteriol* 187:3950–3959
- Free A, Dorman CJ (1995) Coupling of *Escherichia coli hns* mRNA levels to DNA synthesis by autoregulation: implications for growth phase control. *Mol Microbiol* 18:101–113
- Free A, Dorman CJ (1997) The *Escherichia coli stpA* gene is transiently expressed during growth in rich medium and is induced in minimal medium and by stress conditions. *J Bacteriol* 179:909–918
- Free A, Williams RM, Dorman CJ (1998) The StpA protein functions as a molecular adapter to mediate repression of the *bgl* operon by truncated H-NS in *Escherichia coli*. *J Bacteriol* 180:994–997
- Free A, Porter ME, Deighan P, Dorman CJ (2001) Requirement for the molecular adapter function of StpA at the *Escherichia coli bgl* promoter depends upon the level of truncated H-NS protein. *Mol Microbiol* 42:903–917
- Friedrich K, Gualerzi CO, Lammi M, Losso MA, Pon CL (1988) Proteins from the prokaryotic nucleoid. Interaction of nucleic acids with the 15 kDa *Escherichia coli* histone-like protein H-NS. *FEBS Lett* 229:197–202
- Gally DL, Bogan JA, Eisenstein BI, Blomfield IC (1993) Environmental regulation of the fim switch controlling type 1 fimbrial phase variation in *Escherichia coli* K-12: effects of temperature and media. *J Bacteriol* 175:6186–6193
- Garcia J, Cordeiro TN, Nieto JM, Pons I, Juarez A, Pons M (2005) Interaction between the bacterial nucleoid associated proteins Hha and H-NS involves a conformational change of Hha. *Biochem J* 388:755–762
- Garcia J, Madrid C, Juarez A, Pons M (2006) New roles for key residues in helices H1 and H2 of the *Escherichia coli* H-NS N-terminal domain: H-NS dimer stabilization and Hha binding. *J Mol Biol* 359:679–689
- Gerstel U, Park C, Romling U (2003) Complex regulation of *csgD* promoter activity by global regulatory proteins. *Mol Microbiol* 49:639–654
- Gil R, Silva FJ, Zientz E, Delmotte F, Gonzalez-Candelas F, Latorre A, Rausell C, Kamerbeek J, Gadau J, Holldobler B, van Ham RC, Gross R, Moya A (2003) The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc Natl Acad Sci USA* 100:9388–9393
- Godessart N, Munoa FJ, Regue M, Juarez A (1988) Chromosomal mutations that increase the production of a plasmid-encoded haemolysin in *Escherichia coli*. *J Gen Microbiol* 134:2779–2787

- Golub E, Bailone A, Devoret R (1988) A gene encoding an SOS inhibitor is present in different conjugative plasmids. *J Bacteriol* 170:4392–4394
- Goransson M, Sonden B, Nilsson P, Dagberg B, Forsman K, Emanuelsson K, Uhlin BE (1990) Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature* 344:682–685
- Gordon BR, Imperial R, Wang L, Navarre WW, Liu J (2008) Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J Bacteriol* 190:7052–7059
- Gowrishankar J, Pittard AJ (1998) Superimposition of *tyrR* protein-mediated regulation on osmo-responsive transcription of *Escherichia coli proU* in vivo. *J Bacteriol* 180:6743–6748
- Goyard S (1996) Identification and characterization of BpH2, a novel histone H1 homolog in *Bordetella pertussis*. *J Bacteriol* 178:3066–3071
- Goyard S, Bertin P (1997) Characterization of BpH3, an H-NS-like protein in *Bordetella pertussis*. *Mol Microbiol* 24:815–823
- Graeme-Cook KA, May G, Bremer E, Higgins CF (1989) Osmotic regulation of porin expression: a role for DNA supercoiling. *Mol Microbiol* 3:1287–1294
- Grainger DC, Hurd D, Goldberg MD, Busby SJ (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res* 34:4642–4652
- Grainger DC, Aiba H, Hurd D, Browning DF, Busby SJ (2007) Transcription factor distribution in *Escherichia coli*: studies with FNR protein. *Nucleic Acids Res* 35:269–278
- Grainger DC, Goldberg MD, Lee DJ, Busby SJ (2008) Selective repression by Fis and H-NS at the *Escherichia coli dps* promoter. *Mol Microbiol* 68:1366–1377
- Gralla JD, Vargas DR (2006) Potassium glutamate as a transcriptional inhibitor during bacterial osmoregulation. *EMBO J* 25:1515–1521
- Groisman EA, Ochman H (1997) How *Salmonella* became a pathogen. *Trends Microbiol* 5:343–349
- Gualerzi CO, Losso MA, Lammi M, Friedrich K, Pawlik RT, Canonaco MA, Gianfranceschi GL, Pingoud A, Pon CL (1986) Proteins from the Prokaryotic Nucleoid. Structural and Functional Characterization of the *Escherichia coli* DNA-Binding Proteins NS (HU) and H-NS. In: Gualerzi CO, Pon CL (eds) *Bacterial chromatin*. Springer-Verlag, Berlin, pp 101–134
- Haack KR, Robinson CL, Miller KJ, Fowlkes JW, Mellies JL (2003) Interaction of Ler at the *LEE5 (tir)* operon of enteropathogenic *Escherichia coli*. *Infect Immun* 71:384–392
- Halsey TA, Vazquez-Torres A, Gravidahl DJ, Fang FC, Libby SJ (2004) The ferritin-like Dps protein is required for *Salmonella enterica* serovar Typhimurium oxidative stress resistance and virulence. *Infect Immun* 72:1155–1158
- Haniford DB (2006) Transpososome dynamics and regulation in Tn10 transposition. *Crit Rev Biochem Mol Biol* 41:407–424
- Hardy CD, Cozzarelli NR (2005) A genetic selection for supercoiling mutants of *Escherichia coli* reveals proteins implicated in chromosome structure. *Mol Microbiol* 57:1636–1652
- Harrison JA, Pickard D, Higgins CF, Khan A, Chatfield SN, Ali T, Dorman CJ, Hormaeche CE, Dougan G (1994) Role of *hns* in the virulence phenotype of pathogenic *salmonellae*. *Mol Microbiol* 13:133–140
- Heroven AK, Dersch P (2006) RovM, a novel LysR-type regulator of the virulence activator gene *rovA*, controls cell invasion, virulence and motility of *Yersinia pseudotuberculosis*. *Mol Microbiol* 62:1469–1483
- Heroven AK, Nagel G, Tran HJ, Parr S, Dersch P (2004) RovA is autoregulated and antagonizes H-NS-mediated silencing of invasins and *rovA* expression in *Yersinia pseudotuberculosis*. *Mol Microbiol* 53:871–888
- Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, May G, Bremer E (1988) A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. *Cell* 52:569–584
- Higgins CF, Hinton JC, Hulton CS, Owen-Hughes T, Pavitt GD, Seirafi A (1990) Protein H1: a role for chromatin structure in the regulation of bacterial gene expression and virulence? *Mol Microbiol* 4:2007–2012

- Hinton JC, Santos DS, Seirafi A, Hulton CS, Pavitt GD, Higgins CF (1992) Expression and mutational analysis of the nucleoid-associated protein H-NS of *Salmonella typhimurium*. *Mol Microbiol* 6:2327–2337
- Hommais F, Krin E, Laurent-Winter C, Soutourina O, Malpertuy A, Le Caer JP, Danchin A, Bertin P (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* 40:20–36
- Hromockyj AE, Tucker SC, Maurelli AT (1992) Temperature regulation of *Shigella* virulence: identification of the repressor gene *virR*, an analogue of *hns*, and partial complementation by tyrosyl transfer RNA (tRNA_{Tyr}). *Mol Microbiol* 6:2113–2124
- Hulbert RR, Taylor RK (2002) Mechanism of ToxT-dependent transcriptional activation at the *Vibrio cholerae* *tcpA* promoter. *J Bacteriol* 184:5533–5544
- Hulton CS, Seirafi A, Hinton JC, Sidebotham JM, Waddell L, Pavitt GD, Owen-Hughes T, Spassky A, Buc H, Higgins CF (1990) Histone-like protein H1 (H-NS), DNA supercoiling, and gene expression in bacteria. *Cell* 63:631–642
- Jacquet M, Cukier-Kahn R, Pla J, Gros F (1971) A thermostable protein factor acting on *in vitro* DNA transcription. *Biochem Biophys Res Commun* 45:1597–1607
- Johansson J, Uhlin BE (1999) Differential protease-mediated turnover of H-NS and StpA revealed by a mutation altering protein stability and stationary-phase survival of *Escherichia coli*. *Proc Natl Acad Sci USA* 96:10776–10781
- Johansson J, Eriksson S, Sonden B, Wai SN, Uhlin BE (2001) Heteromeric interactions among nucleoid-associated bacterial proteins: localization of StpA-stabilizing regions in H-NS of *Escherichia coli*. *J Bacteriol* 183:2343–2347
- Jordi BJ, Higgins CF (2000) The downstream regulatory element of the *proU* operon of *Salmonella typhimurium* inhibits open complex formation by RNA polymerase at a distance. *J Biol Chem* 275:12123–12128
- Jordi BJ, Dagberg B, de Haan LA, Hamers AM, van der Zeijst BA, Gaastra W, Uhlin BE (1992) The positive regulator CfaD overcomes the repression mediated by histone-like protein H-NS (H1) in the CFA/I fimbrial operon of *Escherichia coli*. *EMBO J* 11:2627–2632
- Jordi BJ, Fielder AE, Burns CM, Hinton JC, Dover N, Ussery DW, Higgins CF (1997) DNA binding is not sufficient for H-NS-mediated repression of *proU* expression. *J Biol Chem* 272:12083–12090
- Joshi R, Passner JM, Rohs R, Jain R, Sosinsky A, Crickmore MA, Jacob V, Aggarwal AK, Honig B, Mann RS (2007) Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* 131:530–543
- Kano Y, Yasuzawa K, Tanaka H, Imamoto F (1993) Propagation of phage Mu in IHF-deficient *Escherichia coli* in the absence of the H-NS histone-like protein. *Gene* 126:93–97
- Kaper JB, Nataro JP, Mobley HL (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2:123–140
- Kawula TH, Orndorff PE (1991) Rapid site-specific DNA inversion in *Escherichia coli* mutants lacking the histone like protein H-NS. *J Bacteriol* 173:4116–4123
- Kim MS, Bae SH, Yun SH, Lee HJ, Ji SC, Lee JH, Srivastava P, Lee SH, Chae H, Lee Y, Choi BS, Chatteraj DK, Lim HM (2005) Cnu, a novel *oriC*-binding protein of *Escherichia coli*. *J Bacteriol* 187:6998–7008
- Klemm P (1986) Two regulatory *fim* genes, *fimB* and *fimE*, control the phase variation of type 1 fimbriae in *Escherichia coli*. *EMBO J* 5:1389–1393
- Kocincova D, Singh AK, Beretti JL, Ren H, Euphrasie D, Liu J, Daffe M, Etienne G, Reyat JM (2008) Spontaneous transposition of IS1096 or ISMsm3 leads to glycopeptidolipid overproduction and affects surface properties in *Mycobacterium smegmatis*. *Tuberculosis (Edinb)* 88:390–398
- La Teana A, Brandi A, Falconi M, Spurio R, Pon CL, Gualerzi CO (1991) Identification of a cold shock transcriptional enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS. *Proc Natl Acad Sci USA* 88:10907–10911
- Laaferki MH, Janabi N, Oswald E, Repoila F (2006) Concert of regulators to switch on LEE expression in enterohemorrhagic *Escherichia coli* O157:H7: interplay between Ler, GrIA, H-NS and RpoS. *Int J Med Microbiol* 296:197–210

- Laine B, Sautiere P, Spassky A, Rimsky S (1984) A DNA-binding protein from *E. coli* isolation, characterization and its relationship with proteins H1 and B1. *Biochem Biophys Res Commun* 119:1147–1153
- Lammi M, Paci M, Pon CL, Losso MA, Miano A, Pawlik RT, Gianfranceschi GL, Gualerzi CO (1984) Proteins from the prokaryotic nucleoid: biochemical and ¹H NMR studies on three bacterial histone-like proteins. In: Hubscher U, Spadari S (eds) *Proteins involved in DNA replication*, vol 179. Plenum, New York, pp 467–477
- Lane D, Cavaille J, Chandler M (1994) Induction of the SOS response by IS1 transposase. *J Mol Biol* 242:339–350
- Lane MC, Simms AN, Mobley HL (2007) Complex interplay between type 1 fimbrial expression and flagellum-mediated motility of uropathogenic *Escherichia coli*. *J Bacteriol* 189:5523–5533
- Lang B, Blot N, Bouffartigues E, Buckle M, Geertz M, Gualerzi CO, Mavathur R, Muskhelishvili G, Pon CL, Rimsky S, Stella S, Babu MM, Travers A (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* 35:6330–6337
- Lawrence JG, Ochman H (1997) Amelioration of bacterial genomes: rates of change and exchange. *J Mol Evol* 44:383–397
- Lawrence JG, Ochman H (1998) Molecular archaeology of the *Escherichia coli* genome. *Proc Natl Acad Sci USA* 95:9413–9417
- Lease RA, Cusick ME, Belfort M (1998) Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc Natl Acad Sci USA* 95:12456–12461
- Lejeune P, Danchin A (1990) Mutations in the *bglY* gene increase the frequency of spontaneous deletions in *Escherichia coli* K-12. *Proc Natl Acad Sci USA* 87:360–363
- Levinthal M, Pownder T (1996) *hns*, *rpoS* and *lrp* mutations affect stationary-phase survival at high osmolarity. *Res Microbiol* 147:333–342
- Lim HM, Lewis DE, Lee HJ, Liu M, Adhya S (2003) Effect of varying the supercoiling of DNA on transcription and its regulation. *Biochemistry* 42:10718–10725
- Lithgow JK, Haider F, Roberts IS, Green J (2007) Alternate SlyA and H-NS nucleoprotein complexes control *hlyE* expression in *Escherichia coli* K-12. *Mol Microbiol* 66:685–698
- Liu Q, Richardson CC (1993) Gene 5.5 protein of bacteriophage T7 inhibits the nucleoid protein H-NS of *Escherichia coli*. *Proc Natl Acad Sci USA* 90:1761–1765
- Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M, Hinton JC (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog* 2:e81
- Lucht JM, Dersch P, Kempf B, Bremer E (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled *proU* operon of *Escherichia coli*. *J Biol Chem* 269:6578–6586
- Madhusudan S, Paukner A, Klingen Y, Schnetz K (2005) Independent regulation of H-NS-mediated silencing of the *bgl* operon at two levels: upstream by BglJ and LeuO and downstream by DnaKJ. *Microbiology* 151:3349–3359
- Madrid C, Nieto JM, Juarez A (2002a) Role of the Hha/YmoA family of proteins in the thermoregulation of the expression of virulence factors. *Int J Med Microbiol* 291:425–432
- Madrid C, Nieto JM, Paytubi S, Falconi M, Gualerzi CO, Juarez A (2002b) Temperature- and H-NS-dependent regulation of a plasmid-encoded virulence operon expressing *Escherichia coli* hemolysin. *J Bacteriol* 184:5058–5066
- Madrid C, Balsalobre C, Garcia J, Juarez A (2007a) The novel Hha/YmoA family of nucleoid-associated proteins: use of structural mimicry to modulate the activity of the H-NS family of proteins. *Mol Microbiol* 63:7–14
- Madrid C, Garcia J, Pons M, Juarez A (2007b) Molecular evolution of the H-NS protein: interaction with Hha-like proteins is restricted to enterobacteriaceae. *J Bacteriol* 189:265–268
- Majdalani N, Cuning C, Sledjeski D, Elliott T, Gottesman S (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci USA* 95:12462–12467

- Mangan MW, Lucchini S, Danino V, Croinin TO, Hinton JC, Dorman CJ (2006) The integration host factor (IHF) integrates stationary-phase and virulence gene expression in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* 59:1831–1847
- Manna D, Wang X, Higgins NP (2001) Mu and IS1 transpositions exhibit strong orientation bias at the *Escherichia coli* *bgl* locus. *J Bacteriol* 183:3328–3335
- Marshall C, Labrousse V, Kreimer M, Weichert D, Kolb A, Hengge-Aronis R (1998) Molecular analysis of the regulation of *csiD*, a carbon starvation-inducible gene in *Escherichia coli* that is exclusively dependent on sigma s and requires activation by cAMP-CRP. *J Mol Biol* 276:339–353
- Masi M, Pages JM, Villard C, Pradel E (2005) The *eefABC* multidrug efflux pump operon is repressed by H-NS in *Enterobacter aerogenes*. *J Bacteriol* 187:3894–3897
- Masson C, Menaa F, Pinon-Lataillade G, Frobert Y, Chevillard S, Radicella JP, Sarasin A, Angulo JF (2003) Global genome repair is required to activate KIN17, a UVC-responsive gene involved in DNA replication. *Proc Natl Acad Sci USA* 100:616–621
- May G, Dersch P, Haardt M, Middendorf A, Bremer E (1990) The *osmZ* (*bglY*) gene encodes the DNA-binding protein H-NS (H1a), a component of the *Escherichia coli* K12 nucleoid. *Mol Gen Genet* 224:81–90
- McClelland M, Florea L, Sanderson K, Clifton SW, Parkhill J, Churcher C, Dougan G, Wilson RK, Miller W (2000) Comparison of the *Escherichia coli* K-12 genome with sampled genomes of a *Klebsiella pneumoniae* and three *Salmonella enterica* serovars, Typhimurium, Typhi and Paratyphi. *Nucleic Acids Res* 28:4974–4986
- McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413:852–856
- McFeeters RL, Altieri AS, Cherry S, Tropea JE, Waugh DS, Byrd RA (2007) The high-precision solution structure of *Yersinia* modulating protein YmoA provides insight into interaction with H-NS. *Biochemistry* 46:13975–13982
- Mellies JL, Barron AM, Carmona AM (2007) Enteropathogenic and enterohemorrhagic *Escherichia coli* virulence gene regulation. *Infect Immun* 75:4199–4210
- Mellies JL, Larabee FJ, Zarr MA, Horback KL, Lorenzen E, Mavor D (2008) Ler interdomain linker is essential for anti-silencing in enteropathogenic *Escherichia coli*. *Microbiology* 154:3624–3638
- Mikulskis AV, Cornelis GR (1994) A new class of proteins regulating gene expression in enterobacteria. *Mol Microbiol* 11:77–86
- Mikulskis AV, Delor I, Thi VH, Cornelis GR (1994) Regulation of the *Yersinia enterocolitica* enterotoxin Yst gene. Influence of growth phase, temperature, osmolarity, pH and bacterial host factors. *Mol Microbiol* 14:905–915
- Mojica FJ, Higgins CF (1997) In vivo supercoiling of plasmid and chromosomal DNA in an *Escherichia coli* *hns* mutant. *J Bacteriol* 179:3528–3533
- Mourino M, Madrid C, Balsalobre C, Prenafeta A, Munoa F, Blanco J, Blanco M, Blanco JE, Juarez A (1996) The Hha protein as a modulator of expression of virulence factors in *Escherichia coli*. *Infect Immun* 64:2881–2884
- Mourino M, Balsalobre C, Madrid C, Nieto JM, Prenafeta A, Munoa FJ, Juarez A (1998) Osmolarity modulates the expression of the Hha protein from *Escherichia coli*. *FEMS Microbiol Lett* 160:225–229
- Mukerji M, Mahadevan S (1997) Characterization of the negative elements involved in silencing the *bgl* operon of *Escherichia coli*: possible roles for DNA gyrase, H-NS, and CRP-cAMP in regulation. *Mol Microbiol* 24:617–627
- Müller CM, Dobrindt U, Nagy G, Emody L, Uhlin BE, Hacker J (2006) Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic *Escherichia coli*. *J Bacteriol* 188:5428–5438

- Murphree D, Froehlich B, Scott JR (1997) Transcriptional control of genes encoding CS1 pili: negative regulation by a silencer and positive regulation by Rns. *J Bacteriol* 179:5736–5743
- Nagarajavel V, Madhusudan S, Dole S, Rahmouni AR, Schnetz K (2007) Repression by binding of H-NS within the transcription unit. *J Biol Chem* 282:23622–23630
- Nair S, Finkel SE (2004) Dps protects cells against multiple stresses during stationary phase. *J Bacteriol* 186:4192–4198
- Navarre WW, Halsey TA, Walthers D, Frye J, McClelland M, Potter JL, Kenney LJ, Gunn JS, Fang FC, Libby SJ (2005) Co-regulation of *Salmonella enterica* genes required for virulence and resistance to antimicrobial peptides by SlyA and PhoP/PhoQ. *Mol Microbiol* 56:492–508
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313:236–238
- Navarre WW, McClelland M, Libby SJ, Fang FC (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev* 21:1456–1471
- Nieto JM, Carmona M, Bolland S, Jubete Y, de la Cruz F, Juarez A (1991) The *hha* gene modulates haemolysin expression in *Escherichia coli*. *Mol Microbiol* 5:1285–1293
- Nieto JM, Prenafeta A, Miquelay E, Torrades S, Juarez A (1998) Sequence, identification and effect on conjugation of the *rmoA* gene of plasmid R100–1. *FEMS Microbiol Lett* 169:59–66
- Nieto JM, Madrid C, Prenafeta A, Miquelay E, Balsalobre C, Carrascal M, Juarez A (2000) Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and H-NS. *Mol Gen Genet* 263:349–358
- Nieto JM, Madrid C, Miquelay E, Parra JL, Rodriguez S, Juarez A (2002) Evidence for direct protein-protein interaction between members of the enterobacterial Hha/YmoA and H-NS families of proteins. *J Bacteriol* 184:629–635
- Nishino K, Yamaguchi A (2004) Role of histone-like protein H-NS in multidrug resistance of *Escherichia coli*. *J Bacteriol* 186:1423–1429
- Noom MC, Navarre WW, Oshima T, Wuite GJ, Dame RT (2007) H-NS promotes looped domain formation in the bacterial chromosome. *Curr Biol* 17:R913–R914
- Normand P, Lapierre P, Tisa LS, Gogarten JP, Alloisio N, Bagnarol E, Bassi CA, Berry AM, Bickhart DM, Choise N, Couloux A, Courmoyer B, Cruveiller S, Daubin V, Demange N, Francino MP, Goltsman E, Huang Y, Kopp OR, Labarre L, Lapidus A, Lavire C, Marechal J, Martinez M, Mastrorunzio JE, Mullin BC, Niemann J, Pujic P, Rawnsley T, Rouy Z, Schenowitz C, Sellstedt A, Tavares F, Tomkins JP, Vallenet D, Valverde C, Wall LG, Wang Y, Medigue C, Benson DR (2007) Genome characteristics of facultatively symbiotic *Frankia sp.* strains reflect host range and host plant biogeography. *Genome Res* 17:7–15
- Norte VA, Stapleton MR, Green J (2003) PhoP-responsive expression of the *Salmonella enterica* serovar Typhimurium *slyA* gene. *J Bacteriol* 185:3508–3514
- Nougayrede JP, Fernandes PJ, Donnenberg MS (2003) Adhesion of enteropathogenic *Escherichia coli* to host cells. *Cell Microbiol* 5:359–372
- Nye MB, Taylor RK (2003) *Vibrio cholerae* H-NS domain structure and function with respect to transcriptional repression of ToxR regulon genes reveals differences among H-NS family members. *Mol Microbiol* 50:427–444
- Ó Cróinín T, Dorman CJ (2007) Expression of the Fis protein is sustained in late-exponential- and stationary-phase cultures of *Salmonella enterica* serovar Typhimurium grown in the absence of aeration. *Mol Microbiol* 66:237–251
- O’Gara JP, Dorman CJ (2000) Effects of local transcription and H-NS on inversion of the fim switch of *Escherichia coli*. *Mol Microbiol* 36:457–466
- Ochman H, Davalos LM (2006) The nature and dynamics of bacterial genomes. *Science* 311:1730–1733
- Oh TJ, Jung IL, Kim IG (2001) The *Escherichia coli* SOS gene *sbmC* is regulated by H-NS and RpoS during the SOS induction and stationary growth phase. *Biochem Biophys Res Commun* 288:1052–1058

- Ohta T, Ueguchi C, Mizuno T (1999) *rpoS* function is essential for *bgl* silencing caused by C-terminally truncated H-NS in *Escherichia coli*. J Bacteriol 181:6278–6283
- Olekhovich IN, Kadner RJ (2006) Crucial roles of both flanking sequences in silencing of the *hilA* promoter in *Salmonella enterica*. J Mol Biol 357:373–386
- Olekhovich IN, Kadner RJ (2007) Role of nucleoid-associated proteins Hha and H-NS in expression of *Salmonella enterica* activators HilD, HilC, and RtsA required for cell invasion. J Bacteriol 189:6882–6890
- Olsen A, Arnqvist A, Hammar M, Sukupolvi S, Normark S (1993) The RpoS sigma factor relieves H-NS-mediated transcriptional repression of *csgA*, the subunit gene of fibronectin-binding curli in *Escherichia coli*. Mol Microbiol 7:523–536
- Ono S, Goldberg MD, Olsson T, Esposito D, Hinton JC, Ladbury JE (2005) H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochem J 391:203–213
- Oshima T, Ishikawa S, Kurokawa K, Aiba H, Ogasawara N (2006) *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. DNA Res 13:141–153
- Overdier DG, Csonka LN (1992) A transcriptional silencer downstream of the promoter in the osmotically controlled *proU* operon of *Salmonella typhimurium*. Proc Natl Acad Sci USA 89:3140–3144
- Owen-Hughes TA, Pavitt GD, Santos DS, Sidebotham JM, Hulton CS, Hinton JC, Higgins CF (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. Cell 71:255–265
- Paci M, Pon CL, Canonaco MA, Gualerzi CO (1986) Proteins from the prokaryotic nucleoid. [1H]-NMR studies on *Escherichia coli* histone-like proteins NS and H-NS and their interactions with DNA. In: Gualerzi CO, Pon CL (eds) Bacterial chromatin. Springer-Verlag, Berlin, pp 135–154
- Pan CQ, Finkel SE, Cramton SE, Feng JA, Sigman DS, Johnson RC (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. J Mol Biol 264:675–695
- Paytubi S, Madrid C, Forns N, Nieto JM, Balsalobre C, Uhlin BE, Juarez A (2004) YdgT, the Hha paralogue in *Escherichia coli*, forms heteromeric complexes with H-NS and StpA. Mol Microbiol 54:251–263
- Perez JC, Latifi T, Groisman EA (2008) Overcoming H-NS-mediated transcriptional silencing of horizontally acquired genes by the PhoP and SlyA proteins in *Salmonella enterica*. J Biol Chem 283:10773–10783
- Pons JJ, Rodriguez S, Madrid C, Juarez A, Nieto JM (2004) *In vivo* increase of solubility of over-expressed Hha protein by tandem expression with interacting protein H-NS. Protein Exp Purif 35:293–297
- Porter ME, Dorman CJ (1994) A role for H-NS in the thermo-osmotic regulation of virulence gene expression in *Shigella flexneri*. J Bacteriol 176:4187–4191
- Porter ME, Dorman CJ (1997) Positive regulation of *Shigella flexneri* virulence genes by integration host factor. J Bacteriol 179:6537–6550
- Postow L, Hardy CD, Arsuaga J, Cozzarelli NR (2004) Topological domain structure of the *Escherichia coli* chromosome. Genes Dev 18:1766–1779
- Prosseda G, Fradiani PA, Di Lorenzo M, Falconi M, Micheli G, Casalino M, Nicoletti M, Colonna B (1998) A role for H-NS in the regulation of the *virF* gene of *Shigella* and enteroinvasive *Escherichia coli*. Res Microbiol 149:15–25
- Prosseda G, Falconi M, Giangrossi M, Gualerzi CO, Micheli G, Colonna B (2004) The *virF* promoter in *Shigella*: more than just a curved DNA stretch. Mol Microbiol 51:523–537
- Pul U, Lux B, Wurm R, Wagner R (2008) Effect of upstream curvature and transcription factors H-NS and LRP on the efficiency of *Escherichia coli* rRNA promoters P1 and P2 – a phasing analysis. Microbiology 154:2546–2558
- Rajkumari K, Gowrishankar J (2001) *In vivo* expression from the RpoS-dependent P1 promoter of the osmotically regulated *proU* operon in *Escherichia coli* and *Salmonella enterica* serovar Typhimurium: activation by *rho* and *hns* mutations and by cold stress. J Bacteriol 183:6543–6550

- Renzoni D, Esposito D, Pfuhl M, Hinton JC, Higgins CF, Driscoll PC, Ladbury JE (2001) Structural characterization of the N-terminal oligomerization domain of the bacterial chromatin-structuring protein, H-NS. *J Mol Biol* 306:1127–1137
- Riley M, Abe T, Arnaud MB, Berlyn MK, Blattner FR, Chaudhuri RR, Glasner JD, Horiuchi T, Keseler IM, Kosuge T, Mori H, Perna NT, Plunkett G 3rd, Rudd KE, Serres MH, Thomas GH, Thomson NR, Wishart D, Wanner BL (2006) *Escherichia coli* K-12: a cooperatively developed annotation snapshot – 2005. *Nucleic Acids Res* 34:1–9
- Rimsky S (2004) Structure of the histone-like protein H-NS and its role in regulation and genome superstructure. *Curr Opin Microbiol* 7:109–114
- Rimsky S, Spassky A (1990) Sequence determinants for H1 binding on *Escherichia coli* lac and gal promoters. *Biochemistry* 29:3765–3771
- Rine J (1999) On the mechanism of silencing in *Escherichia coli*. *Proc Natl Acad Sci USA* 96:8309–8311
- Robbe-Saule V, Schaeffer F, Kowarz L, Norel F (1997) Relationships between H-NS, sigma S, SpvR and growth phase in the control of *spvR*, the regulatory gene of the *Salmonella* plasmid virulence operon. *Mol Gen Genet* 256:333–347
- Rocha EP, Danchin A (2002) Base composition bias might result from competition for metabolic resources. *Trends Genet* 18:291–294
- Rodriguez S, Nieto JM, Madrid C, Juarez A (2005) Functional replacement of the oligomerization domain of H-NS by the Hha protein of *Escherichia coli*. *J Bacteriol* 187:5452–5459
- Rouquette C, Serre MC, Lane D (2004) Protective role for H-NS protein in IS1 transposition. *J Bacteriol* 186:2091–2098
- Schechter LM, Jain S, Akbar S, Lee CA (2003) The small nucleoid-binding proteins H-NS, HU, and Fis affect *hila* expression in *Salmonella enterica* serovar Typhimurium. *Infect Immun* 71:5432–5435
- Schneider R, Lurz R, Luder G, Tolksdorf C, Travers A, Muskhelishvili G (2001) An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. *Nucleic Acids Res* 29:5107–5114
- Schnetz K (1995) Silencing of *Escherichia coli* *bgl* promoter by flanking sequence elements. *EMBO J* 14:2545–2550
- Schnetz K (2002) Silencing of the *Escherichia coli* *bgl* operon by RpoS requires Crl. *Microbiology* 148:2573–2578
- Schnetz K (2008) Fine-tuned growth phase control of *dps*, encoding a DNA protection protein, by FIS and H-NS. *Mol Microbiol* 68:1345–1347
- Schnetz K, Wang JC (1996) Silencing of the *Escherichia coli* *bgl* promoter: effects of template supercoiling and cell extracts on promoter activity *in vitro*. *Nucleic Acids Res* 24:2422–2428
- Schroder O, Wagner R (2000) The bacterial DNA-binding protein H-NS represses ribosomal RNA transcription by trapping RNA polymerase in the initiation complex. *J Mol Biol* 298:737–748
- Seputiene V, Suziedelis K, Normark S, Melefors O, Suziedeliene E (2004) Transcriptional analysis of the acid-inducible *asr* gene in enterobacteria. *Res Microbiol* 155:535–542
- Shi X, Bennett GN (1994) Plasmids bearing *hfq* and the *hns*-like gene *stpA* complement *hns* mutants in modulating arginine decarboxylase gene expression in *Escherichia coli*. *J Bacteriol* 176:6769–6775
- Shi Y, Latifi T, Cromie MJ, Groisman EA (2004) Transcriptional control of the antimicrobial peptide resistance *ugtL* gene by the *Salmonella* PhoP and SlyA regulatory proteins. *J Biol Chem* 279:38618–38625
- Shiga Y, Sekine Y, Kano Y, Ohtsubo E (2001) Involvement of H-NS in transpositional recombination mediated by IS1. *J Bacteriol* 183:2476–2484
- Shin M, Song M, Rhee JH, Hong Y, Kim YJ, Seok YJ, Ha KS, Jung SH, Choy HE (2005) DNA looping-mediated repression by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping. *Genes Dev* 19:2388–2398
- Shindo H, Iwaki T, Jeda R, Kurumizaka H, Ueguchi C, Mizuno T, Morikawa S, Nakamura H, Kuboniwa H (1995) Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from *Escherichia coli*. *FEBS Lett* 360:125–131

- Shindo H, Ohnuki A, Ginba H, Katoh E, Ueguchi C, Mizuno T, Yamazaki T (1999) Identification of the DNA binding surface of H-NS protein from *Escherichia coli* by heteronuclear NMR spectroscopy. *FEBS Lett* 455:63–69
- Silphaduang U, Mascarenhas M, Karmali M, Coombes BK (2007) Repression of intracellular virulence factors in *Salmonella* by the Hha and YdgT nucleoid-associated proteins. *J Bacteriol* 189:3669–3673
- Singh RK, Liburd J, Wardle SJ, Haniford DB (2008) The nucleoid binding protein H-NS acts as an anti-channeling factor to favor intermolecular Tn10 transposition and dissemination. *J Mol Biol* 376:950–962
- Sledjeski D, Gottesman S (1995) A small RNA acts as an antisilencer of the H-NS-silenced *rcaA* gene of *Escherichia coli*. *Proc Natl Acad Sci USA* 92:2003–2007
- Smyth CP, Lundback T, Renzoni D, Siligardi G, Beavil R, Layton M, Sidebotham JM, Hinton JC, Driscoll PC, Higgins CF, Ladbury JE (2000) Oligomerization of the chromatin-structuring protein H-NS. *Mol Microbiol* 36:962–972
- Sondén B, Uhlin BE (1996) Coordinated and differential expression of histone-like proteins in *Escherichia coli*: regulation and function of the H-NS analog StpA. *EMBO J* 15:4970–4980
- Sonnenfeld JM, Burns CM, Higgins CF, Hinton JC (2001) The nucleoid-associated protein StpA binds curved DNA, has a greater DNA-binding affinity than H-NS and is present in significant levels in *hns* mutants. *Biochimie* 83:243–249
- Spassky A, Rimsky S, Garreau H, Buc H (1984) H1a, an *E. coli* DNA-binding protein which accumulates in stationary phase, strongly compacts DNA in vitro. *Nucleic Acids Res* 12:5321–5340
- Spears PA, Schauer D, Orndorff PE (1986) Metastable regulation of type 1 piliation in *Escherichia coli* and isolation and characterization of a phenotypically stable mutant. *J Bacteriol* 168:179–185
- Sperandio V, Mellies JL, Delahay RM, Frankel G, Crawford JA, Nguyen W, Kaper JB (2000) Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler. *Mol Microbiol* 38:781–793
- Spurio R, Durrenberger M, Falconi M, La Teana A, Pon CL, Gualerzi CO (1992) Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. *Mol Gen Genet* 231:201–211
- Spurio R, Falconi M, Brandi A, Pon CL, Gualerzi CO (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *EMBO J* 16:1795–1805
- Stapleton MR, Norte VA, Read RC, Green J (2002) Interaction of the *Salmonella typhimurium* transcription and virulence factor SlyA with target DNA and identification of members of the SlyA regulon. *J Biol Chem* 277:17630–17637
- Starcic-Erjavec M, van Putten JP, Gaastra W, Jordi BJ, Grabnar M, Zgur-Bertok D (2003) H-NS and Lrp serve as positive modulators of *traJ* expression from the *Escherichia coli* plasmid pRK100. *Mol Genet Genom* 270:94–102
- Stella S, Spurio R, Falconi M, Pon CL, Gualerzi CO (2005) Nature and mechanism of the in vivo oligomerization of nucleoid protein H-NS. *EMBO J* 24:2896–2905
- Stella S, Falconi M, Lammi M, Gualerzi CO, Pon CL (2006) Environmental control of the in vivo oligomerization of nucleoid protein H-NS. *J Mol Biol* 355:169–174
- Stoebel DM, Free A, Dorman CJ (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* 154:2533–2545
- Stonehouse E, Kovackova G, Taylor RK, Skorupski K (2008) Integration host factor positively regulates virulence gene expression in *Vibrio cholerae*. *J Bacteriol* 190:4736–4748
- Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrenner P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrook-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, Olson MV (2000) Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature* 406:959–964

- Suzuki T, Ueguchi C, Mizuno T (1996) H-NS regulates OmpF expression through *micF* antisense RNA in *Escherichia coli*. J Bacteriol 178:3650–3653
- Swingle B, O'Carroll M, Haniford D, Derbyshire KM (2004) The effect of host-encoded nucleoid proteins on transposition: H-NS influences targeting of both IS903 and Tn10. Mol Microbiol 52:1055–1067
- Tamas I, Klasson L, Canback B, Naslund AK, Eriksson AS, Wernegreen JJ, Sandstrom JP, Moran NA, Andersson SG (2002) 50 million years of genomic stasis in endosymbiotic bacteria. Science 296:2376–2379
- Tendeng C, Bertin PN (2003) H-NS in Gram-negative bacteria: a family of multifaceted proteins. Trends Microbiol 11:511–518
- Tendeng C, Badaut C, Krin E, Gounon P, Ngo S, Danchin A, Rimsky S, Bertin P (2000) Isolation and characterization of *vicH*, encoding a new pleiotropic regulator in *Vibrio cholerae*. J Bacteriol 182:2026–2032
- Tendeng C, Krin E, Soutourina OA, Marin A, Danchin A, Bertin PN (2003a) A novel H-NS-like protein from an antarctic psychrophilic bacterium reveals a crucial role for the N-terminal domain in thermal stability. J Biol Chem 278:18754–18760
- Tendeng C, Soutourina OA, Danchin A, Bertin PN (2003b) MvaT proteins in *Pseudomonas spp.*: a novel class of H-NS-like proteins. Microbiology 149:3047–3050
- Tietze E, Tschäpe H (1994) Temperature-dependent expression of conjugation pili by IncM plasmid-harboring bacteria: identification of plasmid-encoded regulatory functions. J Basic Microbiol 34:105–116
- Timchenko T, Bailone A, Devoret R (1996) Btdc, a mouse protein that binds to curved DNA, can substitute in *Escherichia coli* for H-NS, a bacterial nucleoid protein. EMBO J 15:3986–3992
- Tippner D, Wagner R (1995) Fluorescence analysis of the *Escherichia coli* transcription regulator H-NS reveals two distinguishable complexes dependent on binding to specific or nonspecific DNA sites. J Biol Chem 270:22243–22247
- Tippner D, Afflerbach H, Bradacsek C, Wagner R (1994) Evidence for a regulatory function of the histone-like *Escherichia coli* protein H-NS in ribosomal RNA synthesis. Mol Microbiol 11:589–604
- Tissier A, Kannouche P, Mauffrey P, Allemand I, Frelat G, Devoret R, Angulo JF (1996) Molecular cloning and characterization of the mouse Kin17 gene coding for a Zn-finger protein that preferentially recognizes bent DNA. Genomics 38:238–242
- Torres AG, Lopez-Sanchez GN, Milflores-Flores L, Patel SD, Rojas-Lopez M, Martinez de la Pena CF, Arenas-Hernandez MM, Martinez-Laguna Y (2007) Ler and H-NS, regulators controlling expression of the long polar fimbriae of *Escherichia coli* O157:H7. J Bacteriol 189:5916–5928
- Tran HJ, Heroven AK, Winkler L, Spreter T, Beatrix B, Dersch P (2005) Analysis of RovA, a transcriptional regulator of *Yersinia pseudotuberculosis* virulence that acts through antirepression and direct transcriptional activation. J Biol Chem 280:42423–42432
- Travers A, Schneider R, Muskhelishvili G (2001) DNA supercoiling and transcription in *Escherichia coli*: the FIS connection. Biochimie 83:213–217
- Tupper AE, Owen-Hughes TA, Ussery DW, Santos DS, Ferguson DJ, Sidebotham JM, Hinton JC, Higgins CF (1994) The chromatin-associated protein H-NS alters DNA topology *in vitro*. EMBO J 13:258–268
- Turner EC, Dorman CJ (2007) H-NS antagonism in *Shigella flexneri* by VirB, a virulence gene transcription regulator that is closely related to plasmid partition factors. J Bacteriol 189:3403–3413
- Typas A, Hengge R (2006) Role of the spacer between the -35 and -10 regions in sigmaS promoter selectivity in *Escherichia coli*. Mol Microbiol 59:1037–1051
- Typas A, Becker G, Hengge R (2007) The molecular basis of selective promoter activation by the sigmaS subunit of RNA polymerase. Mol Microbiol 63:1296–1306
- Ueguchi C, Kakeda M, Mizuno T (1993) Autoregulatory expression of the *Escherichia coli* *hns* gene encoding a nucleoid protein: H-NS functions as a repressor of its own transcription. Mol Gen Genet 236:171–178

- Ueguchi C, Suzuki T, Yoshida T, Tanaka K, Mizuno T (1996) Systematic mutational analysis revealing the functional domain organization of *Escherichia coli* nucleoid protein H-NS. *J Mol Biol* 263:149–162
- Ueguchi C, Seto C, Suzuki T, Mizuno T (1997) Clarification of the dimerization domain and its functional significance for the *Escherichia coli* nucleoid protein H-NS. *J Mol Biol* 274:145–151
- Ulanovsky L, Bodner M, Trifonov EN, Choder M (1986) Curved DNA: design, synthesis, and circularization. *Proc Natl Acad Sci USA* 83:862–866
- Umanski T, Rosenshine I, Friedberg D (2002) Thermoregulated expression of virulence genes in enteropathogenic *Escherichia coli*. *Microbiology* 148:2735–2744
- Uyar E, Kurokawa K, Yoshimura M, Ishikawa S, Ogasawara N, Oshima T (2009) Differential binding profiles of StpA in wild-type and *hns* mutant cells: a comparative analysis of cooperative partners by chromatin immunoprecipitation-microarray analysis. *J Bacteriol* 191:2388–2391
- Vallet I, Diggle SP, Stacey RE, Camara M, Ventre I, Lory S, Lazdunski A, Williams P, Filloux A (2004) Biofilm formation in *Pseudomonas aeruginosa*: fimbrial *cup* gene clusters are controlled by the transcriptional regulator MvaT. *J Bacteriol* 186:2880–2890
- van Passel MW, Marri PR, Ochman H (2008) The emergence and fate of horizontally acquired genes in *Escherichia coli*. *PLoS Comput Biol* 4:e1000059
- van Ulsen P, Hillebrand M, Zulianello L, van de Putte P, Goosen N (1996) Integration host factor alleviates the H-NS-mediated repression of the early promoter of bacteriophage Mu. *Mol Microbiol* 21:567–578
- van Ulsen P, Hillebrand M, Kainz M, Collard R, Zulianello L, van de Putte P, Gourse RL, Goosen N (1997a) Function of the C-terminal domain of the alpha subunit of *Escherichia coli* RNA polymerase in basal expression and integration host factor-mediated activation of the early promoter of bacteriophage Mu. *J Bacteriol* 179:530–537
- van Ulsen P, Hillebrand M, Zulianello L, van de Putte P, Goosen N (1997b) The integration host factor-DNA complex upstream of the early promoter of bacteriophage Mu is functionally symmetric. *J Bacteriol* 179:3073–3075
- Varshavsky AJ, Nedospasov SA, Bakayev VV, Bakayeva TG, Georgiev GP (1977) Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Res* 4:2725–2745
- Vivero A, Banos RC, Mariscotti JF, Oliveros JC, Garcia-del Portillo F, Juarez A, Madrid C (2008) Modulation of horizontally acquired genes by the Hha-YdgT proteins in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 190:1152–1156
- Walthers D, Carroll RK, Navarre WW, Libby SJ, Fang FC, Kenney LJ (2007) The response regulator SsrB activates expression of diverse *Salmonella* pathogenicity island 2 promoters and counters silencing by the nucleoid-associated protein H-NS. *Mol Microbiol* 65:477–493
- Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM (1998) Force and velocity measured for single molecules of RNA polymerase. *Science* 282:902–907
- Ward CM, Wardle SJ, Singh RK, Haniford DB (2007) The global regulator H-NS binds to two distinct classes of sites within the Tn10 transpososome to promote transposition. *Mol Microbiol* 64:1000–1013
- Wardle SJ, O'Carroll M, Derbyshire KM, Haniford DB (2005) The global regulator H-NS acts directly on the transpososome to promote Tn10 transposition. *Genes Dev* 19:2224–2235
- Waterman SR, Small PL (2003) Transcriptional expression of *Escherichia coli* glutamate-dependent acid resistance genes *gadA* and *gadBC* in an *hns rpoS* mutant. *J Bacteriol* 185:4644–4647
- Westermarck M, Oscarsson J, Mizunoe Y, Urbonaviciene J, Uhlin BE (2000) Silencing and activation of ClyA cytotoxin expression in *Escherichia coli*. *J Bacteriol* 182:6347–6357
- Westfall LW, Carty NL, Layland N, Kuan P, Colmer-Hamood JA, Hamood AN (2006) *mvaT* mutation modifies the expression of the *Pseudomonas aeruginosa* multidrug efflux operon *mexEF-oprN*. *FEMS Microbiol Lett* 255:247–254
- Whitfield CR, Wardle SJ, Haniford DB (2009) The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic Acids Res* 37:309–321

- Will WR, Frost LS (2006) Characterization of the opposing roles of H-NS and TraJ in transcriptional regulation of the F-plasmid *tra* operon. *J Bacteriol* 188:507–514
- Will WR, Lu J, Frost LS (2004) The role of H-NS in silencing F transfer gene expression during entry into stationary phase. *Mol Microbiol* 54:769–782
- Williams RM, Rimsky S, Buc H (1996) Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. *J Bacteriol* 178:4335–4343
- Williamson HS, Free A (2005) A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS antagonist. *Mol Microbiol* 55:808–827
- Wolf T, Janzen W, Blum C, Schnetz K (2006) Differential dependence of StpA on H-NS in auto-regulation of *stpA* and in regulation of *bgl*. *J Bacteriol* 188:6728–6738
- Wyborn NR, Stapleton MR, Norte VA, Roberts RE, Grafton J, Green J (2004) Regulation of *Escherichia coli* hemolysin E expression by H-NS and *Salmonella* SlyA. *J Bacteriol* 186:1620–1628
- Xia Y, Forsman K, Jass J, Uhlin BE (1998) Oligomeric interaction of the PapB transcriptional regulator with the upstream activating region of pili adhesin gene promoters in *Escherichia coli*. *Mol Microbiol* 30:513–523
- Yamada H, Muramatsu S, Mizuno T (1990) An *Escherichia coli* protein that preferentially binds to sharply curved DNA. *J Biochem (Tokyo)* 108:420–425
- Yamashino T, Ueguchi C, Mizuno T (1995) Quantitative control of the stationary phase-specific sigma factor, sigma S, in *Escherichia coli*: involvement of the nucleoid protein H-NS. *EMBO J* 14:594–602
- Yang J, Tauschek M, Strugnell R, Robins-Browne RM (2005) The H-NS protein represses transcription of the *eltAB* operon, which encodes heat-labile enterotoxin in enterotoxigenic *Escherichia coli*, by binding to regions downstream of the promoter. *Microbiology* 151:1199–1208
- Yarmolinsky M (2000) Transcriptional silencing in bacteria. *Curr Opin Microbiol* 3:138–143
- Yee A, Chang X, Pineda-Lucena A, Wu B, Semesi A, Le B, Ramelot T, Lee GM, Bhattacharyya S, Gutierrez P, Denisov A, Lee CH, Cort JR, Kozlov G, Liao J, Finak G, Chen L, Wishart D, Lee W, McIntosh LP, Gehring K, Kennedy MA, Edwards AM, Arrowsmith CH (2002) An NMR approach to structural proteomics. *Proc Natl Acad Sci USA* 99:1825–1830
- Yu RR, DiRita VJ (2002) Regulation of gene expression in *Vibrio cholerae* by ToxT involves both antirepression and RNA polymerase stimulation. *Mol Microbiol* 43:119–134
- Zhang A, Belfort M (1992) Nucleotide sequence of a newly-identified *Escherichia coli* gene, *stpA*, encoding an H-NS-like protein. *Nucleic Acids Res* 20:6735
- Zhang A, Derbyshire V, Salvo JL, Belfort M (1995) *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly in vitro. *RNA* 1:783–793
- Zhang A, Rimsky S, Reaban ME, Buc H, Belfort M (1996) *Escherichia coli* protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *EMBO J* 15:1340–1349
- Zhang J, Zeuner Y, Kleefeld A, Unden G, Janshoff A (2004) Multiple site-specific binding of Fis protein to *Escherichia coli* *nuoA-N* promoter DNA and its impact on DNA topology visualised by means of scanning force microscopy. *ChemBiochem* 5:1286–1289
- Zhou Y, Gottesman S (2006) Modes of regulation of RpoS by H-NS. *J Bacteriol* 188:7022–7025
- Zuber F, Kotlarz D, Rimsky S, Buc H (1994) Modulated expression of promoters containing upstream curved DNA sequences by the *Escherichia coli* nucleoid protein H-NS. *Mol Microbiol* 12:231–240

Chapter 14

FIS and Nucleoid Dynamics upon Exit from Lag Phase

Georgi Muskhelishvili and Andrew Travers

Abstract FIS is perhaps one of the most spectacular members of the small class of bacterial nucleoid-associated proteins. The transient pattern of FIS expression is unique and for no other protein of this class has been the spatial arrangement of multiple DNA binding sites so conspicuously related to function. The importance of the helical arrangement of FIS binding sites is especially obvious in synaptic complexes of DNA invertases and the transcription initiation complexes of stable RNA promoters. It is now apparent, that FIS can bind DNA cooperatively and that variable arrangements of DNA sites in conjunction with a wide range of binding site discrimination underlie the versatility and dynamics of the FIS nucleoprotein complexes. Furthermore, as an integral component of the homeostatic network regulating cellular DNA topology, FIS is so far the only highly abundant nucleoid-associated protein directly implicated in the control of genes coding for major cellular DNA topoisomerases. Nevertheless, the influence of FIS on global nucleoid architecture remains largely unknown. In this article we argue that modulation of DNA supercoil dynamics by FIS upon exit from lag phase facilitates conformational transitions of the nucleoprotein structures organized at two different – local and global – levels of complexity. At each level, the regulatory device involving FIS converts the analog information provided by supercoil dynamics of DNA into digital information uniquely encoded in the regulated gene(s). On this view, the mechanism of transcriptional regulation by topological transitions in the DNA molecule manifests a fractal character.

G. Muskhelishvili (✉)

Jacobs University, Campus Ring 1, D-28759, Bremen, Germany
e-mail: g.muskhelishvili@jacobs-university.de

A. Travers

MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 0QH, UK
Fondation Pierre-Gilles de Gennes, c/o LBPA, Ecole Normale
Supérieure de Cachan 9423, Cachan, France
e-mail: aat@mrc-lmb.cam.ac.uk

Keywords DNA architectural proteins • DNA transactions • genetic network • growth phase-dependent transcription • nucleoid structure • supercoiling

14.1 Introduction

Factor for Inversion Stimulation (FIS) was initially identified as an accessory protein stimulating the site-specific recombination of DNA catalysed by DNA invertases (Kahmann et al. 1985; Koch and Kahmann 1986; Johnson et al. 1986; Haffter and Bickle 1987). However, following work established that this abundant DNA architectural protein is a pleiotropic regulator modulating not only distinct site-specific DNA recombination reactions (Ball and Johnson 1991) but also the major cellular DNA transactions including chromosomal DNA replication (Thompson et al. 1987; Gille et al. 1991; Filutowicz et al. 1992; Roth et al. 1994; Ryan et al. 2004), stable RNA transcription (Nilsson et al. 1990; Ross et al. 1990; Lazarus and Travers 1993), DNA transposition and illegitimate recombination (Bétermier et al. 1993; van Drunen et al. 1993; Shanado et al. 1997).

FIS production sharply increases on commitment of cells to exponential growth rapidly reaching high concentrations (about 50 μM), sufficient to cover almost a quarter of the bacterial chromosome and declines steeply towards the late exponential phase (Ball et al. 1992). FIS binds as a homodimer to multiple DNA sites with affinities differing by three orders of magnitude and correspondingly bends DNA to varying degrees in vitro (Thompson and Landy 1988; Bétermier et al. 1994; Pan et al. 1996). The FIS binding site consensus represents a degenerate 15 bp “core” sequence normally rich in AT base pairs and flanked by G and C bases at its 5' and 3' ends, respectively. However, the sequences flanking the 15 bp “core” can substantially influence the binding and bending of DNA by FIS (Pan et al. 1996; Shao et al. 2008). In general, FIS prefers bendable DNA sequences (Bétermier et al. 1994), perhaps because FIS binds into adjacent major grooves, while the separation between these latter in the canonical B-form DNA exceeds the distance between recognition helices of the two DNA binding HTH domains of FIS (Kostrewa et al. 1991; Yuan et al. 1991; Merickel et al. 2002). In vitro FIS binding is also sensitive to DNA structural dynamics preferring moderately supercoiled templates (Schneider et al. 1997).

The accumulation of FIS on entry into log phase suggests that at this stage most of the regulation is mediated through direct FIS binding. Indeed, early biochemical and genetic studies identified high-affinity FIS binding sequences in the replication origin, attachment sites of phage, recombinational enhancers and the upstream activating sequences (UASs) of stable RNA transcripts (Gille et al. 1991; Thompson et al. 1987; Johnson et al. 1986; Verbeek et al. 1990; Zacharias et al. 1992; Lazarus and Travers 1993; Perkins-Balding et al. 1997). However, some effects of FIS on DNA transactions could not be explained by specific binding implicating an indirect regulation mechanism (Bétermier et al. 1993; van Drunen et al. 1993; Shanado et al. 1997; Margulies and Kaguni 1998). Three lines of evidence

provided support for this “indirect” mode of FIS action. First, FIS was shown to regulate the expression of DNA topoisomerases and modulate the overall topology of DNA in growing cells (Schneider et al. 1997, 1999; Weinstein-Fischer et al. 2000; Keane and Dorman 2003; Ohniwa et al. 2006; Weinstein-Fischer and Altuvia 2007). Second, binding of FIS to multiple DNA sites *in vitro* was shown to constrain supercoils, stabilize branched plectonemes and DNA loops, suggesting that cooperative binding of FIS affects the long-range structure of DNA (Fig. 14.1; Bétermier et al. 1994; Schneider et al. 1997, 2001; Margulies and Kaguni 1998; Skoko et al. 2006). Finally, FIS was found to stabilize topological domain barriers in the nucleoid *in vivo* (Hardy and Cozzarelli 2005), although the three-dimensional DNA structure stabilized by FIS *in vivo* is less obvious. It is clear however, that the global influence of FIS on cellular physiology cannot be explained solely via the interactions with specific high-affinity binding sites compiled in the transcriptional regulatory network (TRN) of the RegulonDB database (Salgado et al. 2006). While we believe that the same is true for H-NS and perhaps the other nucleoid-associated proteins (NAPs) for which high-affinity binding sites have been identified, it is likely that the specific sites can nucleate cooperative binding of adjacent lower affinity sites and formation of long-range DNA structures (Dame 2005; Lang et al. 2007; Maurer et al. 2009). The stability of such structures depends on the growth phase-dependent composition of NAPs, leading to structural oscillation of regulatory nucleoprotein complexes engaged in DNA transactions during bacterial growth

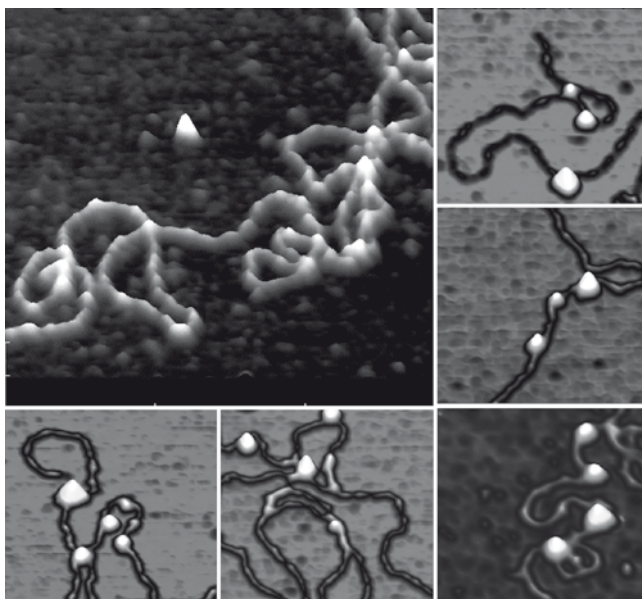


Fig. 14.1 Atomic force microscopy images of FIS nucleoprotein complexes formed with linear 48.5 kb λ DNA. The molar ratio of FIS to DNA is one dimer per 1,165 bps. Note the bending and looping of DNA on binding FIS, which appears as white aggregates associated with DNA. Central panel – three-dimensional rendering of image (Courtesy S. Maurer)

(Gonzalez-Gil et al. 1998; McLeod and Johnson 2001; Nasser et al. 2001; Dorman and Deighan 2003; Ryan et al. 2004; Browning et al. 2004, 2005; Dame 2005). Therefore, the cooperation and competition of NAPs for binding regulatory sequences is a problem of considerable interest.

14.2 The Overarching Network Regulating DNA Superhelicity

Understanding of the role of FIS in the control of cellular physiology requires the definition of a proper context in which the protein operates. During the last decade it became increasingly apparent that *fis* is connected in an overarching homeostatic network regulating overall DNA superhelicity in the cell. This network comprises the NAPs, DNA topoisomerases and also the components of transcription machinery (Bensaid et al. 1996; Schneider et al. 1997, 1999, 2000; Travers et al. 2001; Hirsch and Elliott 2005; Travers and Muskhelishvili 2005a; Blot et al. 2006; Lang et al. 2007; Ó Cróinín and Dorman 2007; Muskhelishvili and Travers 2009). Global DNA superhelicity is determined by opposing activities of DNA gyrase and the DNA relaxing topoisomerases I and IV (Menzel and Gellert 1983; Zechiedrich et al. 1997, 2000) and responds instantly to changes in growth conditions (Drlica 1992; Dorman 1996; Tse-Dinh et al. 1997), thus coordinating genomic expression and cellular metabolism (Cheung et al. 2003; Peter et al. 2004; Blot et al. 2006). The NAPs, and in particular FIS and H-NS, are implicated in modulating the spatiotemporal distributions of free DNA supercoils available to the transcription machinery in the genome (Travers and Muskhelishvili 2005a; Blot et al. 2006; Muskhelishvili and Travers 2009). Unsurprisingly, most if not all the genes in this homeostatic network respond themselves to changes of superhelicity. In addition, the expression of network component genes is interdependent, such that this network has a “heterarchical” rather than hierarchical structure (Fig. 14.2). The cross-talk between the network components is especially conspicuous at the level of transcriptional control and FIS was found to regulate expression of several other NAPs, DNA topoisomerases and components of transcription machinery including the RNA polymerase (RNAP) RpoS (σ^S) and RpoZ (ω) subunits (Schneider et al. 1999; Keane and Dorman 2003; Weinschtein-Fischer and Altuda 2007; Hirsch and Elliott 2005; Geertz et al. 2009). In turn, *fis* expression is subject to negative autoregulation and is also regulated by several components of the network including IHF, CRP, H-NS, RpoS and DksA (Ball et al. 1992; Ninnemann et al. 1992; Pratt et al. 1997; Gonzalez-Gil et al. 1998; Mallik et al. 2006; Ó Cróinín and Dorman 2007; Lang et al. 2007). Furthermore, expression of FIS is sensitive to metabolic state directly responding to variations of both, the nutritional richness, DNA negative superhelicity and also small molecules including nucleosidetriphosphates (NTPs) and ppGpp, the effector molecule of “stringent response” (Nilsson et al. 1992a,b; Ninnemann et al. 1992; Schneider et al. 2000; Walker et al. 2004).

The regulatory effects exerted by FIS upon exit from lag phase normally occur in the context of increased overall negative superhelicity and predominance of the vegetative form of RNAP ($E\sigma^{70}$), as well as those NAPs abundant in early exponential

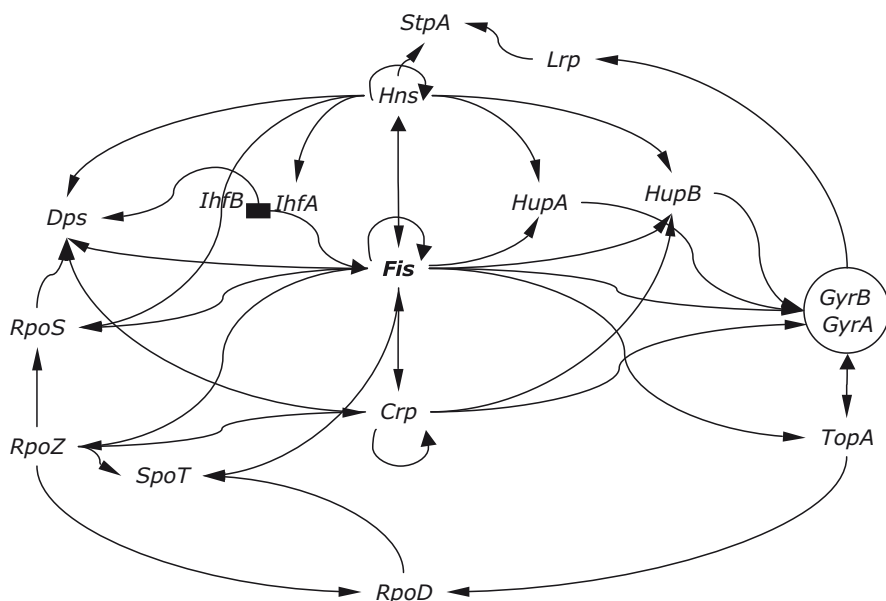


Fig. 14.2 The homeostatic overarching regulatory network of NAPs, topoisomerases and transcription machinery components. The organization of the network is compiled from available literature and unpublished data. Despite being incomplete the network clearly features a heterarchical structure, by virtue of which any component alteration is transmitted to the entire network responding as a whole

phase, e.g. HU α 2 and H-NS (Fig. 14.3; Balke and Gralla 1987; Travers and Muskhelishvili 2005a; Muskhelishvili and Travers 2009). At this growth stage high concentrations of FIS also repress *rpoS* transcription thus reducing the σ^S levels (Hirsch and Elliott 2005). However, both the observations with isolated genes and studies of the global transcript profiles indicate that the regulatory effect of FIS is not confined to exponential growth but is observed also at later stages including the stationary phase, when the impact of E σ^S holoenzyme is maximal and the amount of FIS drops to undetectable levels (Xu and Johnson 1995a,b; Blot et al. 2006; Bradley et al. 2007; Grainger et al. 2008). Since the high initial concentration of FIS is gradually diluted by subsequent cellular division cycles, it has been proposed that the temporal gradient of FIS expression is determinative for the dynamic alterations of chromatin architecture and transcription during the entire growth cycle (Schneider et al. 1997).

Thus, it appears that understanding the mechanism of coordination of the nucleoid architecture and transcription by FIS requires an integration of the systemic aspect determined by the organization of homeostatic network imposing constraints on the activity of its components, and the structural aspect determined by competitive and cooperative interactions of FIS with other NAPs, RNAP and DNA topoisomerases modulating the shape of DNA. Notably, the first aspect is more related to processing of information within the heterarchical network, whereas the other is pertinent to conversion of this information into three-dimensional structure of DNA. However, since the heterarchical

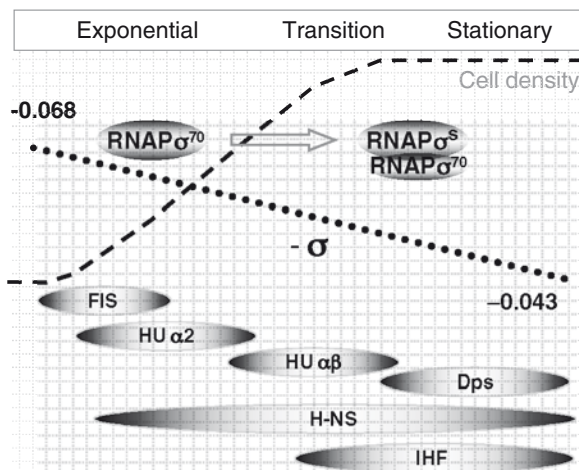


Fig. 14.3 Cellular DNA topology, the composition of NAPs and transcription machinery components change with growth phase. The abundance of particular NAPs and forms of RNAP holoenzyme is approximated for the exponential growth, transition state and stationary phase. The apparent decrease of superhelical density ($-\sigma$) is indicated by dotted line

network operates in the context of the bacterial nucleoid, such that at any instant the processing of information is reflected in a coordinated assembly and disassembly of genomic nucleoprotein complexes, at first approximation this network will have its counterpart in corresponding overall nucleoid structure (Stuger et al. 2002). Therefore, integration of the systemic and structural aspects of regulation by FIS can be attempted by analyzing the dynamic properties of nucleoprotein assemblies stabilized by FIS at different levels of organizational complexity.

14.3 FIS as a Component of Dynamic Nucleoprotein Complexes

The transcriptional regulation of cellular genes at the bottom line involves concerted untwisting of the genomic promoters enabling coordinated gene expression. This necessarily entails a topological transition in the DNA molecule. Indeed, harnessing of superhelical energy for local DNA melting is assumed to be a general mechanism utilized by nucleoprotein complexes involved in regulation of transcription initiation (Travers and Muskhelishvili 1998, 2007; Hatfield and Benham 2002). By the same token, superhelical energy is utilized by the other major DNA transactions requiring local untwisting of DNA sites, including site-specific recombination, replication and DNA transposition (Skarstad et al. 1990; Lim and Simon 1992; Muskhelishvili and Travers 1997; Chalmers et al. 1998; Kondo et al. 2000). Many of these nucleoprotein complexes, at least transiently, contain FIS as a dynamic architectural element.

14.3.1 *OriC* Prereplication Complexes

High amounts of FIS produced during early exponential phase boost ribosome production to support rapid growth (Nilsson et al. 1992a), and therefore it is perhaps not surprising that FIS is also involved in regulation of chromosomal DNA replication. While the expression of *fis* gene is under both the growth phase and growth rate control (Ninnemann et al. 1992; Ball et al. 1992; Nilsson et al. 1992b), it is assumed that under conditions of rapid growth FIS regulates the precision of timing and synchrony of DNA replication initiation (Ryan et al. 2004). Interestingly, the measurements of the *oriC* to *Ter* ratios in the populations of wild type and *fis* mutant cells did not reveal any noticeable differences, whereas analysis of single cells shows that the distribution of the *oriC* to *Ter* ratios in the *fis* mutant are much wider, that in the wild type (Fig. 14.4). This is in keeping with the asynchrony of replication initiation and impaired maintenance of minichromosomes observed in *fis* mutant cells (Gille et al. 1991; Filutowicz et al. 1992). Furthermore, the *fis* mutation

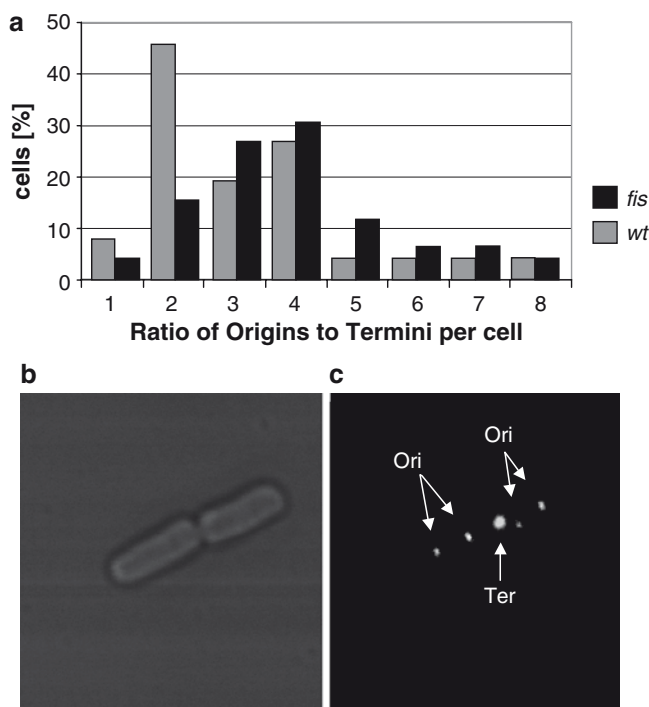


Fig. 14.4 The *oriC* to *Ter* ratio is perturbed in *fis* mutant cells. (a) Distributions of *oriC*/*Ter* ratios determined for individual cells. Note that in wild type cells the highest observed ratios are 2 and 4 as expected, whereas the *fis* mutant demonstrates a wider distribution. (b) and (c) – respectively the phase contrast and fluorescence microscopy images of wild type cells carrying distinct arrays of operators inserted close to *oriC* and to *Ter* and visualised by binding of corresponding fluorescent repressor proteins (Lau et al. 2003) (Courtesy M. Berger)

acts synergistically with *gyrB* mutations in suppressing *oriC*-dependent replication at high temperatures (Filutowicz et al. 1992), consistent with the crosstalk between FIS and gyrase in supercoiling homeostasis (Schneider et al. 2000; Keane and Dorman 2003). Indeed, the initiation of *oriC* replication was demonstrated to depend on negative supercoiling suggesting a directional propagation of DNA untwisting towards the origin (Skarstad et al. 1990; von Freiesleben and Rasmussen 1992), and *oriC*-dependent plasmid replication was shown to critically depend on the superhelical density constrained by FIS (Margulies and Kaguni 1998). These observations implicate FIS in regulating dynamic topological transitions of *oriC* DNA. FIS was shown to specifically bind *oriC* competing with binding of the initiator protein DnaA (Gille et al. 1991). However, since *oriC* also binds IHF the dynamic occupation of high-affinity sites by FIS and IHF appears pivotal for their corresponding effects on *oriC* DNA topology and replication initiation (Roth et al. 1994; Ryan et al. 2004). Importantly, while FIS competes for binding with both IHF and HU in vitro and inhibits *oriC* replication initiation to various extents depending on the composition and structure of the nucleoprotein complex (Wold et al. 1996), the relief of repression by FIS is associated with an IHF-dependent redistribution of DnaA binding and abrupt conformational transition from repressed complex to a completely loaded unwound complex (Ryan et al. 2004). It was suggested that this redistribution of DnaA binding to multiple DnaA boxes constrains a right-handed superhelix (Erzberger et al. 2006), and that a compensatory increase of negative writhe in the adjacent region facilitates the untwisting of the *oriC* DNA. It is noteworthy, that the constraint of positive superhelicity by DnaA protein can be due either to right-handed toroidal wrapping of DNA (Erzberger et al. 2006), or to reducing the negative writhe of a plectoneme, whereas binding of FIS stabilises negatively supercoiled plectonemes and left-handed toroidal coils with low pitch (Schneider et al. 2001; Maurer et al. 2006). This suggests that the structure of the FIS-*oriC* DNA nucleoprotein complex counteracts the topological transitions induced by IHF-dependent redistribution of DnaA binding. By contrast HU, which constrains a high pitch toroid, stimulates strand opening at *oriC* in vitro (Hwang and Kornberg 1992; Bahloul et al. 2001; Chodavarapu et al. 2008). Thus, FIS appears as one of the dynamic architectural proteins modulating the structure of prereplication complexes in which the conformational transitions driven by superhelical energy promote local untwisting of *oriC* DNA.

14.3.2 *Synaptic Complexes of DNA Invertases*

For site-specific DNA inversion reaction to occur, two dimers of DNA invertase bound to recombination sites have to form a tetrameric enzyme by bridging two DNA duplexes in a dynamic nucleoprotein structure termed the synaptic complex, or the invertasome. In the invertasome FIS fulfills a role of both an architectural element stabilising a branched DNA structure and an activator of DNA untwisting and strand cleavage by the invertase tetramer (Heichman and Johnson 1990;

Kanaar et al. 1990; Merickel et al. 1998; Nanassy and Hughes 1998). In addition to FIS, at least in some in vitro recombination systems the DNA inversion is facilitated by HU (Paull et al. 1994). However, in all in vitro systems the assembly of a productive invertasome requires negative supercoiling and the frequency of DNA inversion also critically depends on the level of negative superhelicity (Benjamin et al. 1996; Lim and Simon 1992). Binding of FIS bends DNA at two high-affinity sites in the so-called “recombinational enhancer” sequence, which becomes intertwined with the crossover sites in a branched DNA structure of a productive invertasome. The FIS-enhancer subcomplex acts as a “topological filter” facilitating the trapping of two negative supercoils during invertasome assembly (Crisona et al. 1994; Merickel and Johnson 2004), whereas the recombination reaction is associated with dissipation of the stored negative superhelicity (Kanaar et al. 1990; Benjamin et al. 1996; Merickel and Johnson 2004). Importantly, while the distance between the recombinational enhancer and the crossover sites is largely irrelevant, the helical arrangement of FIS binding sites is critical for enhancer function, suggesting a requirement for specific DNA geometry induced on bending of the enhancer DNA by FIS (Johnson et al. 1987; Perkins-Balding et al. 1997). Furthermore, the tight intertwining of the crossover sites with the FIS-enhancer complex in the invertasome suggests that transitions in the geometry of the enhancer complex can be coordinated with the conformational alterations of invertase tetramer to facilitate the untwisting of crossover sites and coordinated strand cleavage (Haykinson et al. 1996; Perkins-Balding et al. 1997). While the strand cleavage is stimulated by protein interactions between the flexible N-terminal domain of FIS and the dimer interface of the invertase (Osuna et al. 1991; Koch et al. 1991; Merickel et al. 1998), the untwisting of crossover sites appears critically dependent on the flexibility of invertase contact interfaces (Haykinson et al. 1996; Lee et al. 2001). Indeed, the isolated FIS-independent mutants of the invertase Gin with an altered dimerisation interface show enhanced untwisting of the crossover sites in the absence of FIS (Klippel et al. 1988, 1993). After aiding into the assembly of the productive invertasome and triggering the recombination reaction the FIS-enhancer sub-assembly can be released from the Gin synaptic complex (Kanaar et al. 1990). Thus here again, FIS acts as a dynamic architectural element of synaptic complexes facilitating the conformational transitions and promoting local untwisting of recombination sites and strand transfer (Benjamin et al. 1996).

14.3.3 Transcription Initiation Complexes of Stable RNA Promoters

The transcriptional control of stable RNA synthesis maintains the balance between ribosome production and nutrient availability and is therefore central for bacterial growth (Travers and Muskhelishvili 2005a). The ribosomal RNA operons possess exceptionally strong promoters and appear as extended structures densely packed by polymerase with nascent transcripts forming Christmas tree patterns in chromatin

spreads in vitro (French and Miller 1989). In vivo these operons organize local accumulations of RNAP in the nucleoid – the transcription foci (Cabrera and Jin 2003). On transition of the cells to log phase FIS acts as major activator of the stable RNA promoters, which require high negative superhelical density for optimal transcription (Oostra et al. 1981; Ohlsen and Gralla 1992; Bowater et al. 1994; Free and Dorman 1994).

Investigations of the molecular mechanism of activation of stable RNA promoters by FIS provided a paradigm regarding the organization of the ternary initiation complex of the $E\sigma^{70}$ holoenzyme (Maurer et al. 2006). In the upstream activating sequences (UASs) of the stable RNA promoters multiple FIS binding sites are phased in helical register (Verbeek et al. 1992; Zacharias et al. 1992; Finkel and Johnson 1993; Lazarus and Travers 1993). Their role in transcriptional activation provoked two distinct models, one of which argued that the promoter-proximal FIS binding site I is both, necessary and sufficient for full activation (Gosink et al. 1993). This model was supported by observations that the flexible loop of FIS bound at site I interacts with the α -CTD of polymerase, very much alike CRP, which normally binds a single upstream site activating the promoter (Lavigne et al. 1992; Bokal et al. 1995, 1997; Aiyar et al. 2002). In an alternative model binding of FIS to helically phased sites was proposed to induce coherent bending of UAS and stabilization of a DNA microloop trapping the polymerase at the promoter (Lazarus and Travers 1993; Muskhelishvili et al. 1995). This model was based on the observation that in the *tyrT* promoter UAS the average separation between three FIS binding sites is 10.2–10.3 bp suggesting that at full occupancy of binding sites FIS would stabilize a left-handed toroidal loop (Lazarus and Travers 1993). Indeed, disruption of helical phasing of FIS binding sites impairs both the trapping and activating effects of *tyrT* UAS (Muskhelishvili et al. 1995, 1997). Furthermore, binding of FIS at the far upstream site of the *rrnA*P1 promoter UAS was shown to rescue the promoter activity on DNA relaxation by constraining an additional negative supercoil (Rochman et al. 2004). These data clearly indicate the importance of the local geometry of UAS DNA as inferred also for the recombinational enhancer. The impact of UAS is most pronounced under limitations of RNAP binding, unfavorable salt concentrations or sub-optimal supercoiling regime leading to proposal that UAS confers selective advantage under conditions unfavorable for stable RNA transcription (Lazarus and Travers 1993; Muskhelishvili et al. 1995, 1997). Accordingly, at the *tyrT* promoter the activating effect of FIS increases on deviations from optimal superhelicity in vitro (Auner et al. 2003). A clear dependence of the FIS effect on negative superhelicity has been observed also at the *rrnA* P1 and *leuV* promoters (Rochman et al. 2002; Opel et al. 2004).

Notably, the UAS DNA itself is anisotropically bendable (Drew and Travers 1984). Consistently, on supercoiled *tyrT* DNA the interaction of RNAP with the upstream sequence comprising the promoter-distal FIS binding site III in UAS is enhanced by two orders of magnitude (Pemberton et al. 2002). In keeping with this observation, in the binary transcription initiation complex visualized by AFM about 150 base pairs of DNA are wrapped by polymerase, as opposed to 60 bps wrapped in the absence of UAS. Binding of FIS to three helically phased sites in UAS wraps about 80 bp of DNA as a left-handed toroidal coil forming a sub-complex tightly

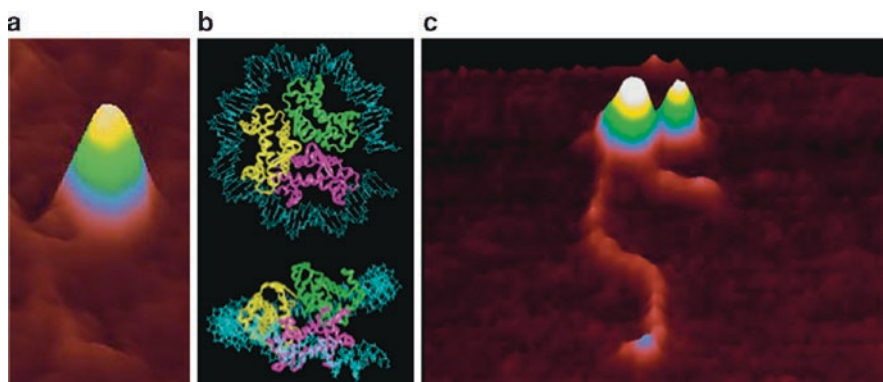


Fig. 14.5 Atomic force microscopy images of nucleoprotein complexes formed with linear *tyrT* promoter DNA. (a) Binding of FIS alone to *tyrT* UAS constrains a right-handed toroid. (b) Model of the FIS – UAS DNA complex in which three FIS dimers binding at helically phased sites in the *tyrT* UAS are assumed to coherently bend the DNA (courtesy T. Hermann). (c) Binding of FIS and RNAP stabilize a ternary complex wrapping over 150 bp of DNA (Reproduced from Maurer et al. 2006. With permission)

attached to RNAP in the ternary complex of FIS, RNAP and *tyrT* promoter DNA (Fig. 14.5; Maurer et al. 2006). Like the invertasome, the ternary initiation complex is highly dynamic in solution (Muskhelishvili et al. 1995, 1997). Indeed, the cooperative binding of FIS to helically phased sites in *tyrT* UAS accelerates ternary complex formation, but in the absence of initiating NTPs a rapid decay of the formed complexes is observed (Muskhelishvili et al. 1997). Since the stable RNA promoters initiate optimally once every second on average, the FIS-dependent decay of the ternary complex in the absence of NTPs might reflect the need of removing stalled initiation complexes to allow rapid re-initiation. Further evidence for the dynamism of the FIS – *tyrT* UAS nucleoprotein complex was obtained by time-resolved photo-crosslinking showing a sequential vacation and re-occupation of the FIS binding site II during ternary complex formation (Pemberton et al. 2002). Such a sequential vacation – re-occupation of DNA binding site has been implicated in utilisation of supercoil energy in a completely different system – Tn10 transpososome, where IHF binding, dissociation and rebinding was associated with supercoil release and the transpososome assembly (Chalmers et al. 1998).

The investigations on the model system of the *tyrT* promoter led to the proposal of the “torsional transmission” model of transcription activation (Muskhelishvili et al. 1997; Muskhelishvili and Travers 1997; Travers and Muskhelishvili 1998). In this model the initiation of transcription is regulated by transition in DNA geometry driven by negative superheicity and facilitated by FIS within a topologically closed domain constrained by the ternary complex. This transition converts the free energy of negative writhe into localized untwisting of the promoter DNA. The AFM structure of a ternary complex formed by FIS and RNAP at the *tyrT* promoter wholly supports this model, implicating in torsional transmission a relative rotation of the two sub-complexes with respect to each other. This rotational movement within a

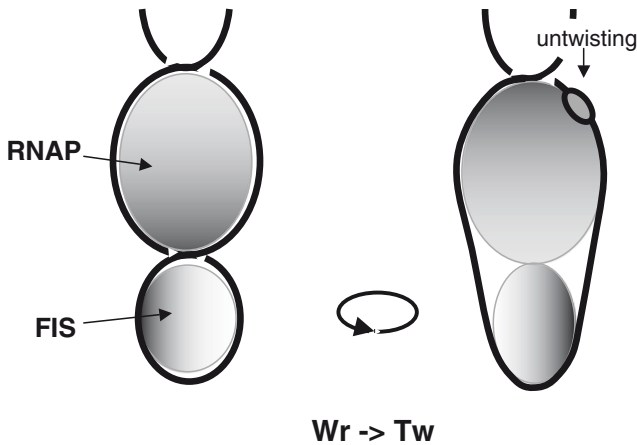


Fig. 14.6 Transcriptional activation by torsional transmission. Schematic of the ternary complex containing stable RNA promoter DNA (black line), RNAP and the FIS-UAS subcomplex stabilizing negative writhe (*left panel*). The DNA torque induced in a closed topological domain by rotation of the FIS-UAS subcomplex with respect to RNAP converts the node of negative writhe into the untwisting of DNA strands at the promoter start site (*right panel*)

topologically closed domain can repartition negative writhe into the untwisting of the -10 element (Fig. 14.6). During initiation, this topological transition in the ternary complex can be facilitated by protein contacts between FIS and both the C-terminal domain of the α subunit (α CTD) and potentially also the σ^{70} subunit of RNAP (Muskhelishvili et al. 1995; Bokal et al. 1997). It is noteworthy however, that at the *leuV* operon promoter FIS binding is proposed to redirect a pre-existent DNA distortion from upstream region to the promoter-proximal region and facilitate promoter opening even in the absence of protein contacts between FIS and the α CTD (Opel et al. 2004). The structure of FIS complex bound to the UAS region of stable RNA promoters also has the potential to facilitate promoter escape. When the transcribing polymerase moves away from the promoter negative superhelicity will likely be generated immediately upstream (Liu and Wang 1987). The ability of FIS to wrap a negative supercoil could in principle enable escape by acting as a torsional sink.

Taken together, it appears that FIS plays a similar role in the dynamic nucleoprotein complexes involved in *oriC* replication, site-specific recombination and stable RNA transcription. In all these DNA transactions, FIS facilitates both the assembly of complex structures and their subsequent conformational transitions channeling the superhelical energy into local untwisting of functionally relevant DNA sites.

14.3.4 Other Promoters

The role of FIS in regulation of global cellular metabolism has been postulated more than a decade ago in experiments comparing the patterns of distinct protein

species in the *E. coli* wild type and *fis* mutant cells (González-Gil et al. 1996). Many genes so identified encoded enzymes or transport proteins involved in the catabolism of sugars and nucleic acids, and their expression was also dependent on the cAMP-CRP complex. For most of these proteins the regulation by FIS occurred to be indirect and mediated through effects on the synthesis of the respective repressor proteins. Indeed, investigations of the transcript profiles of *fis* mutants demonstrate altered expression of many regulatory genes, especially under variations of overall negative superhelicity (Table 1; Blot et al. 2006). It appears therefore that FIS selects the appropriate subsets of regulators (including NAPs), which subsequently interact in the context of compositionally dynamic nucleoprotein complexes regulating the growth phase-dependent gene transcription. Since these regulatory proteins can both cooperate and compete for particular promoter binding, the instant transcriptional effect of such nucleoprotein complexes can be approximated from the growth phase-dependent variations of protein composition in conjunction with spatial organization of the regulator binding sites. Between FIS and CRP, for example, both synergy and competition have been observed dependent on relative arrangement of their cognate DNA binding sites (Nasser et al. 2001; Galán et al. 2008). This variability of interactions facilitates the assembly of temporal regulation patterns sustaining cellular growth.

The growth phase-dependent dynamics of nucleoprotein complexes regulating transcription are obvious from the example of the *fis* promoter itself. It was observed that binding of IHF in the promoter upstream region activates *fis* transcription during transition to exponential growth, whereas at increasing concentrations FIS occupies multiple sites including a site overlapping the core promoter and subsequently lead to autorepression (Ninnemann et al. 1992; Ball et al. 1992; Pratt et al. 1997). Nevertheless, tight repression of the *fis* promoter in late exponential phase requires formation of a nucleoprotein complex in which a CRP molecule is apparently “sandwiched” between two FIS dimers (Nasser et al. 2001). Likewise, the growth phase-dependent expression of *crp* from two tandem promoters is under the control of nucleoprotein complexes of variable composition. During early log phase, binding of FIS at multiple sites overlapping the supercoiling-dependent *crpP2* promoter represses *crp* expression, whereas later on, when FIS level subsides, binding of cAMP-CRP complex induces strong divergent transcription limiting the *crpP1* promoter initiation (Gonzalez-Gil et al. 1998). This mutual repression of *fis* and *crp* expression apparently coordinates the effects of FIS and CRP, which as global regulators share many targets (González-Gil et al. 1996; McLeod et al. 2000). Indeed, while FIS is abundant on commitment of cells to exponential growth in rich medium, CRP is activated by elevated cAMP concentrations on exhaustion of glucose in the growth medium and switches the metabolism to utilization of alternative carbon sources (Travers et al. 2001).

Studies carried out on isolated promoters of important metabolic genes also suggest remarkable structural dynamics of nucleoprotein complexes involving FIS together with more dedicated regulators in optimizing the cellular metabolism during the growth cycle. For example, at the promoter of *ndh* gene encoding the non-proton-translocating NADH dehydrogenase II, FIS represses transcription on binding three

sites centered at -123 , -72 and $+51$ from initiation start in decreasing order of affinity. Notably, only the latter site is repressive, while the two other sites are activating. However, at low FIS concentrations the activation of the *ndh* gene is counteracted by binding of FNR preventing the interactions between FIS and the α CTD of RNAP, whereas at higher concentrations FIS occupies the low-affinity repressor site precluding RNAP binding and open complex formation (Green et al. 1996; Jackson et al. 2004). In contrast, in early exponential phase FIS under respiratory growth conditions activates the transcription of *nuoA-N* operon encoding the proton-pumping NADH dehydrogenase I (Wackwitz et al. 1999). This coordinated effect suggests that FIS favors the utilization of the energetically efficient proton-translocating NADH dehydrogenase I instead of the non-proton translocating NADH dehydrogenase II, thus ensuring higher yields of ATP under rapid growth conditions.

Another well-studied example is the *E. coli nir* operon encoding the NADH-dependent nitrite reductase. Expression of this operon is coordinated by the availability of oxygen, nitrate and nitrite. The *nir* promoter can be fully activated by FNR binding a class II activator site centered at position -41.5 , but this activation is modulated by binding of IHF and FIS to the upstream region (Wu et al. 1998; Browning et al. 2000). It was observed that while FIS, IHF and FNR can bind simultaneously at the *nir* promoter, both FIS and IHF can act as repressors interfering with RNA polymerase binding (Browning et al. 2004; Browning et al. 2008). Again, the observed effects on transcription vary depending on site occupation. Transcription initiation is repressed by binding of FIS and IHF to DNA sites centered at the positions -142 and -88 respectively, whereas IHF binding to a site centered at position -115 activates transcription. Binding of IHF at this latter site apparently facilitates FNR activation by relieving the repressive effect mediated by binding of FIS and IHF at the two other sites. In addition, the regulation of *nir* operon promoter involves also the dedicated transcription factors NarL and NarP, which counteract repression and by remodeling the nucleoprotein structure facilitate the FNR-dependent activation (Browning et al. 2004).

While in most of the studied examples activation of transcription by FIS involves binding sites located in promoter upstream regions, on transition to stationary phase activation of transcription by FIS involves specific binding to a class II activator site centered at position -41 , a location implicated in effects of dedicated transcriptional regulators interacting with σ^{70} subunit of RNAP (Xu and Johnson 1997; McLeod et al. 1999). This observation suggests that different strategies of binding site organization may optimise the effect of FIS on transcription during exponential growth and transition to stationary phase, when the concentration of FIS is high and low, respectively.

These examples in conjunction with the observed altered expression of regulatory genes in *fis* mutant cells (Table 14.1), make it apparent that FIS both acts as a transcriptional regulator of the global and dedicated transcription factors (see Table 14.1), and also interacts with these very same factors in the context of regulatory nucleoprotein complexes. This illustrates a general principle of optimizing the transcriptional control of metabolic function by selecting the regulators subsequently engaged in compositionally distinct nucleoprotein complexes responding to changing growth conditions. For all those nucleoprotein complexes that at least transiently contain FIS, the major question is as to how their organization is coordinated in the genome.

Table 14.1 Distribution of up-regulated transcriptional regulator proteins

	A		B		C		D	
	CSH50	CSH50 <i>fis</i>	Rel wt	Rel <i>fis</i>	Hyp	Rel	Hyp <i>fis</i>	Rel <i>fis</i>
ex	H-NS	Fnr	FruR	AllR	FIS	HU α	LRP	IHF β
	HU α	AllR	EvgA	NarL	ArgR	ArsR	AlpA	Fnr
	BolA	ArgR	ZraR	OxyR	NanR	AsnC	AscG	ArsR
	SoxS	HcaR	GcvA	PaaX	ExuR	CynR	CdaR	GadE
	XylR	OgrK	RfaH	Crl	CreB	EvgA	CpxR	GadW
	FucR	PaaX	SspA	PutA	Crl	IscR	CytR	GadX
	KdgR	Ttk			FlhD	NhaR	Crl	GltF
		DgoR			GcvA	OgrK	CsgD	NhaR
ts	–	–	Hyp wt	Hyp <i>fis</i>	LeuO	OxyR	CusR	SspA
	DicA				LrhA	SspA	DeoR	TrpR
	EnvY		MarA	StpA	NarP	TrpR	EvgA	
	GadE	CitB	MarR	AllR	RcsB		ExuR	
	GadW	QseB	PurR	NarL	PhoP		FlhCD	
	UvrY	RhaS		OxyR	PurR		FruR	
	XapR			Rob	SlyA		HdtR	
	OgrK			NadR	RfaH		HipB	
				TdcA	XylR		FrlR	
st	–	–		ArgR			MalT	
	SlyA	ExuR		Cbl			MalY	
	Ttk	LrhA		SoxS			SlyA	
	CspE	SgrR		HdfR			PurR	
		UxuR		MalY			NadR	
		CreB					NanR	
		DhaR					PspF	
		PspA					QseB	
							Zur	

A – Expression profiles of growing CSH50 wild-type and *fis* mutant cells during exponential phase, transition state and stationary phase (ex, ts and st, respectively). B – Expression profiles comparing exponentially growing wild type and *fis* cells under conditions of relaxation (Rel) and hypernegative superhelicity (Hyp). C – Expression profiles comparing wild type cells under conditions of relaxation and hypernegative superhelicity. D – Expression profiles comparing *fis* cells under conditions of relaxation and hypernegative superhelicity (data compiled from Blot et al. 2006)

14.4 Reorganisation of the Nucleoid Structure by FIS

Comparative DNA microarray analyses of the transcript profiles in wild type and *fis* mutant cells demonstrated similar groups of metabolic genes affected in *Salmonella* and *E. coli* (Kelly et al. 2004; Blot et al. 2006; Bradley et al. 2007). These studies reveal FIS as a systemic regulator directly or indirectly influencing the expression of hundreds of metabolic genes at the level of transcriptional control. Important functions regulated by FIS in *E. coli* and *Salmonella typhimurium* include motility and chemotaxis, energy metabolism, transport, stress response and cell division, whereby in the latter organism FIS is also proposed to coordinate the expression of the functions of motility and adherence with virulence (Kelly et al.

2004; Baxter and Jones 2005; Bradley et al. 2007). Also in the plant pathogen *Erwinia chrysanthemi* FIS was shown to coordinate the entire pathogenicity program, from the production of virulence factors to their translocation into external medium (Lautier et al. 2007). Recent observation of a decreased invasiveness of *Salmonella* *fis* mutant was associated with a counterbalance between the effects of FIS and σ^S in optimizing the transcriptional regulatory network under conditions required for invasion (Ó Cróinín and Dorman 2007). This optimisation also involves a crosstalk between FIS and negative superhelicity facilitating the adaptation of *Salmonella* by providing a mechanism for distinguishing the intra- and extra-cellular conditions (Ó Cróinín et al. 2006). Observed interdependence of FIS and supercoiling in optimizing the physiological response to environmental changes is fully consistent with both the FIS-dependent buffering of stable RNA promoters on deviations of negative superhelicity (Rochman et al. 2002; Auner et al. 2003) and the distinct transcriptional responses of the *E. coli* wild type and mutant cells to directional changes of supercoiling (Blot et al. 2006).

The tight interdependence of FIS and supercoiling in organizing cellular transcription implicates FIS in coupling of the nucleoid architecture and cellular metabolism. Several lines of evidence support the notion that FIS exerts a global effect on nucleoid architecture. In addition to alterations of cellular topoisomerase activities, global superhelical density and the distribution of topological domain barriers in the genome (Schneider et al. 1997; Keane and Dorman 2003; Hardy and Cozzarelli 2005), FIS can counteract the condensation of the nucleoid induced by overexpression of Dps, an abundant NAP which normally compacts the DNA in stationary phase (Ohniwa et al. 2006). The computational analysis of the distribution of FIS binding motif in the *E. coli* genome indicates enrichment for putative FIS binding sites around the terminus of chromosomal replication (Ussery et al. 2001). A genome-wide non-random binding of FIS was suggested from microscopic observations of distinct accumulations of FIS in the nucleoid supported by Chip-on-chip experiments showing accumulation of FIS mainly in the intergenic regions in vivo, as expected for a DNA architectural protein (Azam et al. 2000; Grainger et al. 2006). However, FIS binding was also observed to spread over large distances comprising entire operons in the genome (Grainger et al. 2006). The same is true for binding of H-NS, which predominantly acts as a repressor silencing genes by bridging two DNA duplexes (Dame et al. 2000; Dorman 2004). In contrast to FIS stabilizing loops and branched DNA structures in vitro (Schneider et al. 2001; Skoko et al. 2006), H-NS stabilizes tightly interwound plectonemes (Fig. 14.7; Schneider et al. 2001; Lang et al. 2007; Maurer et al. 2009). Notably, both FIS and H-NS are found at the sites of active transcription, although their ratio may vary depending on the distance from the transcription initiation site as observed after the induction of the *lac* operon (Grainger et al. 2006). Such differential spreading of FIS and H-NS binding over extensive regions in conjunction with various three-dimensional DNA structures can potentially generate topological barriers isolating domains with distinct supercoil dynamics and responses to alterations of global DNA superhelicity (Hardy and Cozzarelli 2005; Blot et al. 2006).

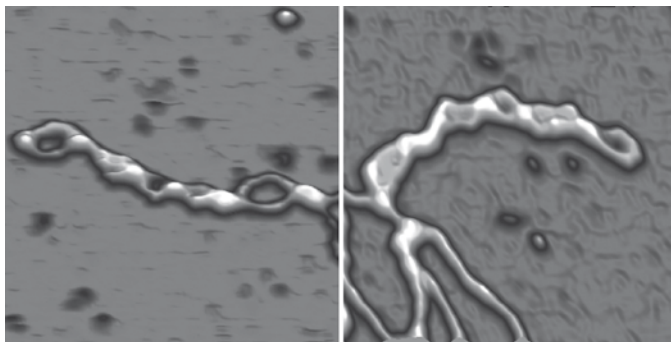


Fig. 14.7 Atomic force microscopy images of H-NS nucleoprotein complexes formed with linear 48.5 kb λ DNA. The molar ratio of H-NS to DNA is one molecule per 1,539 bps. Binding H-NS leads to intertwining of DNA duplicates and formation of a plectonemic filament (Courtesy S. Maurer)

Several studies proposed that the topological domains have variable barriers but are on average 10 kb in extent (Postow et al. 2004; Deng et al. 2005). However, domains of spatially organized transcription of hundreds of kilobases in extent have also been observed and correlated with genomic distributions of gyrase binding sites (Jeong et al. 2004). The impact of FIS on supercoiling sensitivity and spatial organization of genomic transcription was dissected in a study using genetically engineered strains containing drug-resistant topoisomerase gene alleles which enable to selectively inhibit either DNA gyrase or topoisomerase IV activity and thus induce either strong relaxation or high negative superhelicity of genomic DNA by a topoisomerase inhibitor (Khodursky et al. 1995; Zechiedrich et al. 1997). Comparisons of the supercoiling responses of transcription in these strains and isogenic *fis* mutant derivatives distinguished the genes transcribed on induction of relaxation (*Rel*) or high negative supercoiling (*Hyp*) of DNA in exponentially growing cells. Remarkably, although only 10% of genes are common among the total of about 1,500 *Hyp* and *Rel* genes identified in wild type and the *fis* mutant together (Blot et al. 2006), the frequency distributions of the supercoiling responses in the genomes reveal closely coinciding clusters of *Hyp* and *Rel* genes expressed coherently either at high negative superhelicity or at DNA relaxation independent of the genetic background. However, there are also clusters showing deviation between the wild type and mutant. These latter appear to occupy regions adjacent to predicted clusters of FIS binding sites in the genome (Fig. 14.8a). While these findings corroborate previous observations of large transcription domains of variable size and chromosomal macrodomains with variable DNA dynamics (Jeong et al. 2004; Valens et al. 2004; Espeli et al. 2008), they also suggest that deletion of *fis* causes reorganisation of domains of supercoiling-sensitive transcription extending over hundreds of kilobases in the genome.

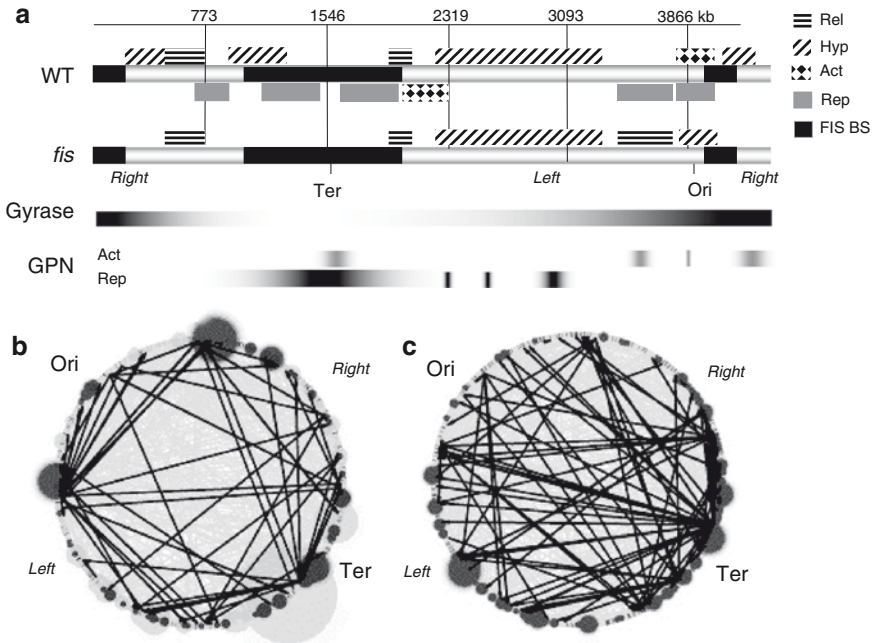


Fig. 14.8 Organisation of the nucleoid structure and supercoiling-sensitive transcription by FIS. (a) Linear representation of the *E. coli* genome sketching the distributions of supercoiling-sensitive domains, FIS binding sites, gyrase binding sites and FIS-dependent GPNs. The genomic positions and the differences in domain organization between the wild-type (WT) and *fis* mutant cells are indicated (courtesy M. Geertz). (b) Genomic distribution of the activating FIS-dependent GPNs (dark grey spheres) and repressing H-NS-dependent GPNs (light-grey spheres) in the circular *E. coli* genome. (c) Genomic distribution of the repressing FIS-dependent GPNs (dark grey spheres) (Reproduced from Marr et al. 2008. With permission.) The connections between the genomic loci in (b) and (c) indicate the nodes of TRNs constrained in wild-type and *fis* mutant cells respectively. Note that the *fis* mutant TRN is more highly connected and that often the FIS-dependent GPNs are closely associated with highly connected TRN nodes. Hyp and Rel, domains of activated transcription at high negative superhelicity and DNA relaxation, respectively; Act and Rep, domains containing genes activated and repressed by FIS. Left and right indicate the replicorees

14.5 Impact of FIS in Distinct Logical Types of Transcriptional Control

How is this organisation of supercoiling-sensitive domains by FIS coupled to the activity of the TRN and the metabolic profile of cells? The impact of the TRN in given transcript pattern can be evaluated by mapping the experimentally obtained expression profiles on the electronically compiled TRN of *E. coli* in the RegulonDB database (Salgado et al. 2006). This database includes all the available information on the experimentally detected regulatory interactions in the organism and features FIS as a “hub” connected to many regulated genes. It is noteworthy

that the information compiled in the TRN is “digital”, as it uses the logic of directional pair-wise connections between a regulator and regulated gene. The mapping procedure enables to estimate the ratio of connected to unconnected nodes in the experimentally obtained transcript profile thus providing an “effective” TRN (the significance of an effective TRN is evaluated by comparisons to corresponding null models comprising random samples with equal numbers of genes). An assessment of the impact of TRN in the transcript profiles of wild type and *fis* mutant cells demonstrated that in mutant the connectivity of the transcriptional regulatory network is significantly higher, than in wild type (Fig. 14.8b; Marr et al. 2008). Such apparently unexpected increase of TRN connectivity on inactivating a highly connected hub is explained by fact that the transcriptional control by high concentrations of FIS produced on entry into the log phase essentially involves differential distributions of superhelical energy and stabilization of supercoiling-sensitive domains in the genome (Fig. 14.8; Travers and Muskhelishvili 2005a,b; Blot et al. 2006). This type of regulation by organizing domains of coherent transcription is dubbed “analog” control, because instead of employing the “digital” information of TRN on the connections between unique genes, it converts a continuous or “analog” property of the genome – distributions of DNA superhelical tension – into the digital pattern of regulated genes. The impact of FIS in analog control can be assessed along the chromosome by measuring the distances between the coherently expressed genes in the transcript profiles. The expressed genes are scored as connected if they fit into a set threshold value. Obtained networks named gene proximity networks (GPNs) are assumed to reflect patterns of structural organisation in the chromosome (Marr et al. 2008). In such analyses FIS was found to constrain both activating and repressing GPNs, the latter being clustered around the *Ter* region (Fig. 14.8), and indicating a substantial loss of spatial connectivity of transcription in mutant cells lacking FIS. Therefore, it is thought that the increased TRN connectivity observed in *fis* mutant is compensatory and represents a novel example of genetic flexibility in bacterial adaptation (Marr et al. 2008). These observations immediately raise the question on the mechanism of such compensatory rearrangement of transcriptional regulation in the *fis* mutant.

Insights into the relationships of distinct types of transcriptional control by FIS are provided by identification of NAPs and dedicated transcriptional regulators shared by the *fis*-dependent GPNs and TRNs. Many GPNs constrained by FIS either contain themselves or are closely associated with one or more of the transcriptional regulator genes (unpublished data). Some of these transcriptional regulator genes also distinctly respond to alterations of supercoiling. Such associations of the transcriptional regulator genes with GPNs suggest how FIS by organizing the supercoiling-sensitive chromosomal domains can coordinate the nucleoid structure with metabolic function, especially since metabolic pathways are distinctly responding to alterations of DNA superhelicity (Blot et al. 2006). In this scenario the increased TRN connectivity observed in the *fis* mutant can be largely explained by derepression of transcriptional regulatory genes associated with the repressing GPNs, which are especially enriched around the *Ter* region. Notably, this region is also enriched in FIS binding sites (Fig. 14.8a,c).

14.6 The Role of FIS in the Overarching Homeostatic Network

The overarching homeostatic network involving FIS in regulating global superhelicity can be considered a self-organising system (Travers and Muskhelishvili 2005a; Karsenti 2008). The sensing mechanism operates in this network at the level of compositional variation of chromosomal nucleoprotein complexes regulating the genomic distributions of effective superhelicity and generating domains of coherent transcription according to the physiological state (Travers and Muskhelishvili 2005a). Such nucleoprotein complexes formed at the stable RNA and other supercoiling-dependent promoters of genes required on transition to lag phase contain FIS acting mainly as an activator. Indeed, molecular studies strongly suggest that activation by FIS is associated with stabilization of local supercoils facilitating conformational transitions and directional channeling of superhelical tension towards the transcription initiation site. This peculiarity makes the effect of FIS strongly dependent on the available superhelical density, which itself is modulated by FIS but ultimately is coordinated within the overarching homeostatic network (see Fig. 14.2). Due to the transient nature of expression, FIS undergoes dynamic interactions with other components of the network including NAPs, DNA topoisomerases and RNA polymerase and thus on transition to lag phase can impose directionality on the system by stabilizing energy-rich toroidal supercoils in the UAS regions of the stable RNA and other gene promoters sustaining rapid growth. Since the RNAP molecules transcribing stable RNA operons are organized in spatially confined foci in the nucleoids of rapidly growing cells (Cabrera and Jin 2003) formation of activating FIS-dependent GPNs would be expected at this stage. Indeed, two major activating FIS GPNs are organized on both sides of the origin, as also are the stable RNA operons, whereas repressing GPNs predominate around the terminus of replication (Fig. 14.8b, c).

Thus it appears that the effect of FIS on nucleoid architecture involves organisation of both, the GPNs spanning several kilobases and activating or repressing transcription of closely spaced genes and operons in the genome (Marr et al. 2008), and also large domains of supercoiling-sensitive transcription extending over hundreds of kilobases (Fig. 14.8). It should be emphasized, that while these structural patterns induced by FIS depend on overall superhelicity, this latter itself is determined by the metabolic state of the cell (McClellan et al. 1990; Hsieh et al. 1991; van Workum et al. 1996; Snoep et al. 2002). Since these supercoiling-dependent structural patterns also engage genes of dedicated transcriptional regulators, the information on the superhelical state of DNA can be directly transmitted to the TRN and ultimately back to the metabolism (Fig. 14.9). Thus, notwithstanding the homeostatic nature of the overarching network, the FIS-dependent reorganization of transcription on exit from lag phase would induce readjustment of the heterarchical network and correspondingly rewire the hierarchical TRN. The structure of the nucleoid would be thus optimised for transcription sustaining rapid growth (Stuger et al. 2002; Muskhelishvili and Travers 2003; Travers and Muskhelishvili 2005a).

In summary, the compilation of the available data (Blot et al. 2006; Marr et al. 2008; unpublished data) on the predicted FIS binding sites clusters, FIS-dependent

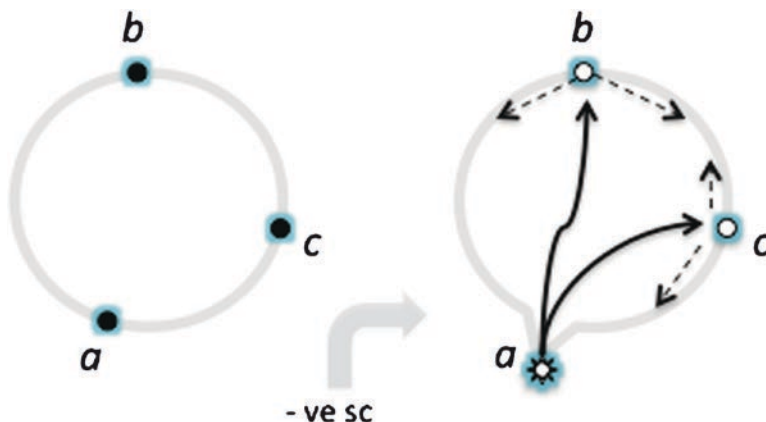


Fig. 14.9 Model of transmission of information on the superhelical state of DNA to the TRN. The circles represent the circular genome. Introduction of superhelical energy ($-ve\ sc$) into the DNA alters the structure of chromatin in the region at *a*, which comprises a transcriptional regulator *a*. Activation of *a* leads to activation of *b* and *c* (arrows), which act as transcriptional regulators of metabolic genes (dashed arrows)

GPNs, and supercoiling-sensitive spatial transcription domains in Fig. 14.8, suggests that FIS coordinates the global structural organisation and transcription of the nucleoid. On the assumption that the state of the heterarchical network can be approximated by overall nucleoid structure, the systemic and structural aspects of regulation by FIS meet at this point. Indeed, the FIS-dependent topological changes of local supercoils regulating the isolated gene promoters and the inferred reorganization of nucleoid structure and global genomic transcription both implicate directional transitions of corresponding three-dimensional DNA structures, although at different levels of organizational complexity. Regarding the flow of genetic information, in both cases FIS facilitates the conversion of analog information provided by supercoil dynamics of DNA into digital information uniquely encoded in specific regulated gene(s) (Blot et al. 2006). This notion of a common generative mechanism implicated in regulation of transcription by FIS at two different levels of organizational complexity reveals a fractal nature of the evolutionary device for channelling genetic information by using controlled topological transitions in the DNA.

References

- Aiyar SE, McLeod SM, Ross W, Hirvonen CA, Thomas MS, Johnson RC, Gourse RL (2002) Architecture of Fis-activated transcription complexes at the *Escherichia coli* *rrnB* P1 and *rrnE* P1 promoters. *J Mol Biol* 316:501–516
- Auner H, Buckle M, Deufel A, Kutateladze T, Lazarus L, Mavathur R, Muskhelishvili G, Pemberton I, Schneider R, Travers A (2003) Mechanism of transcriptional activation by FIS: role of core promoter structure and DNA topology. *J Mol Biol* 331:331–344

- Azam TA, Hiraga S, Ishihama A (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* 5:613–626
- Bahloul A, Boubrik F, Rouvière-Yaniv J (2001) Roles of *Escherichia coli* histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of dnaA46ts. *Biochimie* 83:219–229
- Balke VL, Gralla JD (1987) Changes in the linking number of supercoiled DNA accompany growth transitions in *Escherichia coli*. *J Bacteriol* 169:4499–4506
- Ball CA, Johnson RC (1991) Efficient excision of phage lambda from the *Escherichia coli* chromosome requires the Fis protein. *J Bacteriol* 173:4027–4031
- Ball CA, Osuna R, Ferguson KC, Johnson RC (1992) Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J Bacteriol* 174:8043–8056
- Baxter MA, Jones BD (2005) The *fimYZ* genes regulate *Salmonella enterica* serovar Typhimurium invasion in addition to type 1 fimbrial expression and bacterial motility. *Infect Immun* 73:1377–1385
- Benjamin KR, Abola AP, Kanaar R, Cozzarelli NR (1996) Contributions of supercoiling to Tn3 resolvase and phage Mu Gin site-specific recombination. *J Mol Biol* 256:50–65
- Bensaid A, Almeida A, Drlica K, Rouvière-Yaniv J (1996) Cross-talk between topoisomerase I and HU in *Escherichia coli*. *J Mol Biol* 256:292–300
- Bétermier M, Poquet I, Alazard R, Chandler M (1993) Involvement of *Escherichia coli* FIS protein in maintenance of bacteriophage Mu lysogeny by the repressor: control of early transcription and inhibition of transposition. *J Bacteriol* 175:3798–3811
- Bétermier M, Galas DJ, Chandler M (1994) Interaction of Fis protein with DNA: bending and specificity of binding. *Biochimie* 76:958–967
- Blot N, Mavathur R, Geertz M, Travers A, Muskhelishvili G (2006) Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep* 7:710–715
- Bokal AJ 4th, Ross W, Gourse RL (1995) The transcriptional activator protein FIS: DNA interactions and cooperative interactions with RNA polymerase at the *Escherichia coli* *rrnB* P1 promoter. *J Mol Biol* 245:197–207
- Bokal AJ, Ross W, Gaal T, Johnson RC, Gourse RL (1997) Molecular anatomy of a transcription activation patch: FIS-RNA polymerase interactions at the *Escherichia coli* *rrnB* P1 promoter. *EMBO J* 16:154–162
- Bowater RP, Chen D, Lilley DM (1994) Modulation of *tyrT* promoter activity by template supercoiling *in vivo*. *EMBO J* 13:5647–5655
- Bradley MD, Beach MB, de Koning AP, Pratt TS, Osuna R (2007) Effects of Fis on *Escherichia coli* gene expression during different growth stages. *Microbiology* 153:2922–2940
- Browning DF, Cole JA, Busby SJ (2000) Suppression of FNR-dependent transcription activation at the *Escherichia coli* *nir* promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. *Mol Microbiol* 37:1258–1269
- Browning DF, Cole JA, Busby SJ (2004) Transcription activation by remodelling of a nucleoprotein assembly: the role of NarL at the FNR-dependent *Escherichia coli* *nir* promoter. *Mol Microbiol* 53:203–215
- Browning DF, Grainger DC, Beatty CM, Wolfe AJ, Cole JA, Busby SJ (2005) Integration of three signals at the *Escherichia coli* *nrf* promoter: a role for Fis protein in catabolite repression. *Mol Microbiol* 57:496–510
- Browning DF, Cole JA, Busby SJ (2008) Regulation by nucleoid-associated proteins at the *Escherichia coli* *nir* operon promoter. *J Bacteriol* 190:7258–7267
- Cabrera JE, Jin DJ (2003) The distribution of RNA polymerase in *Escherichia coli* is dynamic and sensitive to environmental cues. *Mol Microbiol* 50:1493–1505
- Chalmers R, Guhathakurta A, Benjamin H, Kleckner N (1998) IHF modulation of Tn10 transposition: sensory transduction of supercoiling status via a proposed protein/DNA molecular spring. *Cell* 93:897–908
- Cheung KJ, Badarinarayana V, Selinger DW, Janse D, Church GM (2003) A microarray-based antibiotic screen identifies a regulatory role for supercoiling in the osmotic stress response of *Escherichia coli*. *Genome Res* 13:206–215

- Chodavarapu S, Felczak MM, Yaniv JR, Kaguni JM (2008) *Escherichia coli* DnaA interacts with HU in initiation at the *E. coli* replication origin. *Mol Microbiol* 67:781–792
- Crisona NJ, Kanaar R, Gonzalez TN, Zechiedrich EL, Klippel A, Cozzarelli NR (1994) Processive recombination by wild-type *gin* and an enhancer-independent mutant. Insight into the mechanisms of recombination selectivity and strand exchange. *J Mol Biol* 243:437–457
- Dame RT (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* 56:858–870
- Dame RT, Wyman C, Goosen N (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* 28:3504–3510
- Deng S, Stein RA, Higgins NP (2005) Organization of supercoil domains and their reorganization by transcription. *Mol Microbiol* 57:1511–1521
- Dorman CJ (1996) Flexible response: DNA supercoiling, transcription and bacterial adaptation to environmental stress. *Trends Microbiol* 4:214–216
- Dorman CJ (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* 2:391–400
- Dorman CJ, Deighan P (2003) Regulation of gene expression by histone-like proteins in bacteria. *Curr Opin Gen Dev* 13:179–184
- Drew HR, Travers AA (1984) DNA structural variations in the *E. coli tyrT* promoter. *Cell* 37:491–502
- Drlica K (1992) Control of DNA supercoiling. *Mol Microbiol* 6:425–433
- Erzberger JP, Mott ML, Berger JM (2006) Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat Struct Mol Biol* 13:676–683
- Espeli O, Mercier R, Boccard F (2008) DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. *Mol Microbiol* 68:1418–1427
- Filutowicz M, Ross W, Wild J, Gourse RL (1992) Involvement of Fis protein in replication of the *Escherichia coli* chromosome. *J Bacteriol* 174:398–407
- Finkel SE, Johnson RC (1993) The Fis protein: its not just for DNA inversion anymore. *Mol Microbiol* 7:1023–1026
- Free A, Dorman CJ (1994) *Escherichia coli tyrT* gene transcription is sensitive to DNA supercoiling in its native chromosomal context: effect of DNA topoisomerase IV overexpression on *tyrT* promoter function. *Mol Microbiol* 14:151–161
- French SL, Miller JR (1989) Transcription mapping of the *Escherichia coli* chromosome by electron microscopy. *J Bacteriol* 171:4207–4216
- Galán B, Manso I, Kolb A, García JL, Prieto MA (2008) The role of FIS protein in the physiological control of the expression of the *Escherichia coli meta-hpa* operon. *Microbiology* 154:2151–2160
- Geertz M, Travers A, Janga SC, Mehandeziska S, Shimamoto N, Muskhelishvili G (2009) Coordination of genomic transcription by bacterial RNA polymerase: role of the omega subunit (submitted)
- Gille H, Egan JB, Roth A, Messer W (1991) The FIS protein binds and bends the origin of chromosomal DNA replication, *oriC*, of *Escherichia coli*. *Nucleic Acids Res* 19:4167–4172
- Gonzalez-Gil G, Kahmann R, Muskhelishvili G (1998) Regulation of *crp* transcription by oscillation between distinct nucleoprotein complexes. *EMBO J* 17:2877–2885
- González-Gil G, Bringmann P, Kahmann R (1996) FIS is a regulator of metabolism in *Escherichia coli*. *Mol Microbiol* 22:21–29
- Gosink KK, Ross W, Leirmo S, Osuna R, Finkel SE, Johnson RC, Gourse RL (1993) DNA binding and bending are necessary but not sufficient for Fis-dependent activation of *rrnB* P1. *J Bacteriol* 175:1580–1589
- Grainger DC, Hurd D, Goldberg MD, Busby S (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res* 34:4642–4652
- Grainger DC, Goldberg MD, Lee DJ, Busby SJ (2008) Selective repression by Fis and H-NS at the *Escherichia coli dps* promoter. *Mol Microbiol* 68:1366–1377
- Green J, Anjum MF, Guest JR (1996) The *ndh*-binding protein (Nbp) regulates the *ndh* gene of *Escherichia coli* in response to growth phase and is identical to Fis. *Mol Microbiol* 20:1043–1055

- Haffter P, Bickle TA (1987) Purification and DNA-binding properties of FIS and Cin, two proteins required for the bacteriophage P1 site-specific recombination system, cin. *J Mol Biol* 198:579–587
- Hardy CD, Cozzarelli NR (2005) A genetic selection for supercoiling mutants of *Escherichia coli* reveals proteins implicated in chromosome structure. *Mol Microbiol* 57:1636–1652
- Hatfield GW, Benham CJ (2002) DNA topology-mediated control of global gene expression in *Escherichia coli*. *Annu Rev Genet* 36:175–203
- Haykinson MJ, Johnson LM, Soong J, Johnson RC (1996) The Hin dimer interface is critical for Fis-mediated activation of the catalytic steps of site-specific DNA inversion. *Curr Biol* 6:163–177
- Heichman KA, Johnson RC (1990) The Hin invertasome: protein-mediated joining of distant recombination sites at the enhancer. *Science* 249:511–517
- Hirsch M, Elliott T (2005) Fis regulates transcriptional induction of RpoS in *Salmonella enterica*. *J Bacteriol* 187:1568–1580
- Hsieh LS, Burger RM, Drlica K (1991) Bacterial DNA supercoiling and [ATP]/[ADP]. Changes associated with a transition to anaerobic growth. *J Mol Biol* 219:443–450
- Hwang DS, Kornberg A (1992) Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF. *J Biol Chem* 267:23083–23086
- Jackson L, Blake T, Green J (2004) Regulation of *ndh* expression in *Escherichia coli* by Fis. *Microbiology* 150:407–413
- Jeong KS, Ahn J, Khodursky AB (2004) Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol* 5:R86
- Johnson RC, Bruist MF, Simon MI (1986) Host protein requirements for *in vitro* site-specific DNA inversion. *Cell* 46:531–539
- Johnson RC, Glasgow AC, Simon MI (1987) Spatial relationship of the Fis binding sites for Hin recombinational enhancer activity. *Nature* 329:462–465
- Kahmann R, Rudt F, Koch C, Mertens G (1985) G inversion in bacteriophage Mu DNA is stimulated by a site within the invertase gene and a host factor. *Cell* 41:771–780
- Kanaar R, Klippel A, Shekhtman E, Dungan JM, Kahmann R, Cozzarelli NR (1990) Processive recombination by the phage Mu Gin system: implications for the mechanisms of DNA strand exchange, DNA site alignment, and enhancer action. *Cell* 62:353–366
- Karsenti E (2008) Self-organization in cell biology: a brief history. *Nat Rev Mol Cell Biol* 9:255–262
- Keane OM, Dorman CJ (2003) The *gyr* genes of *Salmonella enterica* serovar Typhimurium are repressed by the factor for inversion stimulation, Fis. *Mol Gen Genet* 270:56–65
- Kelly A, Goldberg MD, Carroll RK, Danino V, Hinton JC, Dorman CJ (2004) A global role for Fis in the transcriptional control of metabolism and type III secretion in *Salmonella enterica* serovar Typhimurium. *Microbiology* 150:2037–2053
- Khodursky AB, Zechiedrich EL, Cozzarelli NR (1995) Topoisomerase IV is a target of quinolones in *Escherichia coli*. *Proc Natl Acad Sci USA* 92:11801–11805
- Klippel A, Cloppenburg K, Kahmann R (1988) Isolation and characterization of unusual *gin* mutants. *EMBO J* 7:3983–3989
- Klippel A, Kanaar R, Kahmann R, Cozzarelli NR (1993) Analysis of strand exchange and DNA binding of enhancer-independent Gin recombinase mutants. *EMBO J* 12:1047–1057
- Koch C, Kahmann R (1986) Purification and properties of the *Escherichia coli* host factor required for inversion of the G segment in bacteriophage Mu. *J Biol Chem* 261:15673–15678
- Koch C, Ninnemann O, Fuss H, Kahmann R (1991) The N-terminal part of the *E. coli* DNA binding protein FIS is essential for stimulating site-specific DNA inversion but is not required for specific DNA binding. *Nucleic Acids Res* 19:5915–5922
- Kondo T, Mima S, Fukuma N, Sekimizu K, Tsuchiya T, Mizushima T (2000) Suppression of temperature-sensitivity of a *dnaA46* mutant by excessive DNA supercoiling. *Biochem J* 348:375–379

- Kostrewa D, Granzin J, Koch C, Choe HW, Raghunathan S, Wolf W, Labahn J, Kahmann R, Saenger W (1991) Three-dimensional structure of the *E. coli* DNA-binding protein FIS. *Nature* 349:178–180
- Lang B, Blot N, Bouffartigues E, Buckle M, Geertz M, Gualerzi CO, Mavathur R, Muskhelishvili G, Pon CL, Rimsky S, Stella S, Babu MM, Travers A (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* 35:6330–6337
- Lau IF, Filipe SR, Søballe B, Økstad OA, Barre FX, Sherratt DJ (2003) Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol* 49:731–743
- Lautier T, Blot N, Muskhelishvili G, Nasser W (2007) Roles for Fis in the growth phase regulation of pectate lyases, essential virulence factors in the phytopathogenic bacterium *Erwinia chrysanthemi*. *Mol Microbiol* 66:1491–1505
- Lavigne M, Kolb A, Buc H (1992) Transcription activation by cAMP receptor protein (CRP) at the *Escherichia coli galP1* promoter. Crucial role for the spacing between the CRP binding site and the –10 region. *Biochemistry* 31:9647–9656
- Lazarus LR, Travers AA (1993) The *Escherichia coli* FIS protein is not required for the activation of *tyrT* transcription on entry into exponential growth. *EMBO J* 12:2483–2494
- Lee HJ, Lee SY, Lee H, Lim HM (2001) Effects of dimer interface mutations in Hin recombinase on DNA binding and recombination. *Mol Genet Genomics* 266:598–607
- Lim HM, Simon MI (1992) The role of negative supercoiling in Hin-mediated site-specific recombination. *J Biol Chem* 267:11176–11182
- Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci USA* 84:7024–7027
- Mallik P, Paul BJ, Rutherford ST, Gourse RL, Osuna R (2006) DksA is required for growth phase-dependent regulation, growth rate-dependent control, and stringent control of *fis* expression in *Escherichia coli*. *J Bacteriol* 188:5775–5782
- Margulies C, Kaguni JM (1998) The FIS protein fails to block the binding of DnaA protein to *oriC*, the *Escherichia coli* chromosomal origin. *Nucleic Acids Res* 26:5170–5175
- Marr C, Geertz M, Hütt MT, Muskhelishvili G (2008) Dissecting the logical types of network control in gene expression profiles. *BMC Syst Biol* 2:18
- Maurer S, Fritz J, Muskhelishvili G, Travers A (2006) RNA polymerase and an activator form discrete subcomplexes in a transcription initiation complex. *EMBO J* 25:3784–3790
- Maurer S, Fritz J, Muskhelishvili G (2009) A systematic *in vitro* study of nucleoprotein complexes formed by bacterial nucleoid associated proteins revealing novel types of DNA organization. *J Mol Biol* 387:1261–1276
- McClellan JA, Boublíková P, Palecek E, Lilley DMJ (1990) Superhelical torsion in cellular DNA responds directly to environmental and genetic factors. *Proc Natl Acad Sci USA* 87:8373–8377
- McLeod SM, Johnson RC (2001) Control of transcription by nucleoid proteins. *Curr Opin Microbiol* 4:152–159
- McLeod SM, Xu J, Cramton SE, Gaal T, Gourse RL, Johnson RC (1999) Localization of amino acids required for Fis to function as a class II transcriptional activator at the *RpoS*-dependent *proP* P2 promoter. *J Mol Biol* 294:333–346
- McLeod SM, Xu J, Johnson RC (2000) Coactivation of the *RpoS*-dependent *proP* P2 promoter by *fis* and cyclic AMP receptor protein. *J Bacteriol* 182:4180–4187
- Menzel R, Gellert M (1983) Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling. *Cell* 34:105–113
- Merickel SK, Johnson RC (2004) Topological analysis of Hin-catalysed DNA recombination *in vivo* and *in vitro*. *Mol Microbiol* 51:1143–1154
- Merickel SK, Haykinson MJ, Johnson RC (1998) Communication between Hin recombinase and Fis regulatory subunits during coordinate activation of Hin-catalyzed site-specific DNA inversion. *Genes Dev* 12:2803–2816
- Merickel SK, Sanders ER, Vazquez-Ibar JL, Johnson RC (2002) Subunit exchange and the role of dimer flexibility in DNA binding by the Fis protein. *Biochemistry* 41:5788–5798

- Muskhelishvili G, Travers AA (1997) The stabilization of DNA microloops by FIS – a mechanism for torsional transmission in transcription activation and DNA inversion. In: Eckstein F, Lilley DMJ (eds) *Nucleic acids and molecular biology*, Springer 11:179–190
- Muskhelishvili G, Travers A (2003) Transcription factor as a topological homeostat. *Front Biosci* 8:d279–d285
- Muskhelishvili G, Travers A (2009) Intrinsic *in vivo* modulators: negative supercoiling and the constituents of the bacterial nucleoid. In: *RNA polymerases as molecular motors*. R SC publishing, Cambridge, pp 69–95
- Muskhelishvili G, Travers AA, Heumann H, Kahmann R (1995) FIS and RNA polymerase form a specific nucleoprotein complex at a stable RNA promoter. *EMBO J* 14:1446–1452
- Muskhelishvili G, Buckle M, Heumann H, Kahmann R, Travers AA (1997) FIS activates sequential steps during transcription initiation at a stable RNA promoter. *EMBO J* 16:3655–3665
- Nanassy OZ, Hughes KT (1998) In vivo identification of intermediate stages of the DNA inversion reaction catalyzed by the *Salmonella* Hin recombinase. *Genetics* 149:1649–1663
- Nasser W, Schneider R, Travers A, Muskhelishvili G (2001) CRP modulates *fis* transcription by alternate formation of activating and repressing nucleoprotein complexes. *J Biol Chem* 276:17878–17886
- Nilsson L, Vanet A, Vijgenboom E, Bosch L (1990) The role of FIS in trans activation of stable RNA operons of *E. coli*. *EMBO J* 9:727–734
- Nilsson L, Verbeek H, Vijgenboom E, van Druenen C, Vanet A, Bosch L (1992a) FIS-dependent trans activation of stable RNA operons of *Escherichia coli* under various growth conditions. *J Bacteriol* 174:921–929
- Nilsson L, Verbeek H, Hoffmann U, Haupt M, Bosch L (1992b) Inactivation of the *fis* gene leads to reduced growth rate. *FEMS Microbiol Lett* 78:85–88
- Ninnemann O, Koch C, Kahmann R (1992) The *E.coli fis* promoter is subject to stringent control and autoregulation. *EMBO J* 11:1075–1083
- Ó Cróinín T, Dorman CJ (2007) Expression of the Fis protein is sustained in late-exponential- and stationary-phase cultures of *Salmonella enterica* serovar Typhimurium grown in the absence of aeration. *Mol Microbiol* 66:237–251
- Ó Cróinín T, Carroll RK, Kelly A, Dorman CJ (2006) Roles for DNA supercoiling and the Fis protein in modulating expression of virulence genes during intracellular growth of *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* 62:869–882
- Ohlsen KL, Gralla JD (1992) Interrelated effects of DNA supercoiling, ppGpp, and low salt on melting within the *Escherichia coli* ribosomal RNA *rrnB* P1 promoter. *Mol Microbiol* 6:2243–2251
- Ohniwa RL, Morikawa K, Kim J, Ohta T, Ishihama A, Wada C, Takeyasu K (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. *EMBO J* 25:5591–5602
- Oostra BA, van Vliet AJ, Ab G, Gruber M (1981) Enhancement of ribosomal ribonucleic acid synthesis by deoxyribonucleic acid gyrase activity in *Escherichia coli*. *J Bacteriol* 148:782–787
- Opel ML, Aeling KA, Holmes WM, Johnson RC, Benham CJ, Hatfield GW (2004) Activation of transcription initiation from a stable RNA promoter by a Fis protein-mediated DNA structural transmission mechanism. *Mol Microbiol* 53:665–674
- Osuna R, Finkel SE, Johnson RC (1991) Identification of two functional regions in Fis: the N-terminus is required to promote Hin-mediated DNA inversion but not lambda excision. *EMBO J* 10:1593–1603
- Pan CQ, Finkel SE, Cramton SE, Feng JA, Sigman DS, Johnson RC (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *J Mol Biol* 264:675–695
- Paull TT, Haykinson MJ, Johnson RC (1994) HU and functional analogs in eukaryotes promote Hin invertasome assembly. *Biochimie* 76:992–1004
- Pemberton I, Muskhelishvili G, Travers A, Buckle M (2002) FIS modulates the kinetics of successive interactions of RNA polymerase with the core and upstream regions of the *E. coli tyrT* promoter. *J Mol Biol* 318:651–663
- Perkins-Balding D, Dias DP, Glasgow AC (1997) Location, degree and direction of DNA bending associated with the Hin recombinational enhancer sequence and Fis-enhancer complex. *J Bacteriol* 179:4747–4753

- Peter BJ, Arsuaga J, Breier AM, Khodursky AB, Brown PO, Cozzarelli NR (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol* 5:R87
- Postow L, Hardy CD, Arsuaga J, Cozzarelli NR (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev* 18:1766–1779
- Pratt TS, Steiner T, Feldman LS, Walker KA, Osuna R (1997) Deletion analysis of the *fis* promoter region in *Escherichia coli*: antagonistic effects of integration host factor and Fis. *J Bacteriol* 179:6367–6377
- Rochman M, Aviv M, Glaser G, Muskhelishvili G (2002) Promoter protection by a transcription factor acting as a local topological homeostat. *EMBO Rep* 3:355–360
- Rochman M, Blot N, Dyachenko M, Glaser G, Travers A, Muskhelishvili G (2004) Buffering of stable RNA promoter activity against DNA relaxation requires a far upstream sequence. *Mol Microbiol* 53:143–152
- Ross W, Thompson JF, Newlands JT, Gourse RL (1990) *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *EMBO J* 9:3733–3742
- Roth A, Urmoneit B, Messer W (1994) Functions of histone-like proteins in the initiation of DNA replication at *oriC* of *Escherichia coli*. *Biochimie* 76:917–923
- Ryan VT, Grimwade JE, Camara JE, Crooke E, Leonard AC (2004) *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Mol Microbiol* 51:1347–1359
- Salgado H, Gama-Castro S, Peralta-Gil M, Diaz-Peredo E, Sanchez-Solano F, Santos-Zavaleta A, Martinez-Flores I, Jimenez-Jacinto V, Bonavides-Martinez C, Segura-Salazar J et al (2006) RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions. *Nucleic Acids Res* 34:D394–D397
- Schneider R, Travers AA, Muskhelishvili G (1997) FIS regulates the bacterial growth phase-dependent topological transitions in DNA. *Mol Microbiol* 26:519–530
- Schneider R, Travers A, Kutateladze T, Muskhelishvili G (1999) A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Mol Microbiol* 34:953–964
- Schneider R, Travers A, Muskhelishvili G (2000) The expression of the *Escherichia coli fis* gene is strongly dependent on the superhelical density of DNA. *Mol Microbiol* 38:167–175
- Schneider R, Lurz R, Lüder G, Tolksdorf C, Travers A, Muskhelishvili G (2001) An architectural role of the *Escherichia coli* protein FIS in organising DNA. *Nucleic Acids Res* 29:5107–5114
- Shanado Y, Kato JI, Ikeda H (1997) Fis is required for illegitimate recombination during formation of λ bio transducing phage. *J Bacteriol* 179:4239–4245
- Shao Y, Feldman-Cohen LS, Osuna R (2008) Functional characterization of the *Escherichia coli* Fis-DNA binding sequence. *J Mol Biol* 376:771–785
- Skarstad K, Baker TA, Kornberg A (1990) Strand separation required for initiation of replication at the chromosomal origin of *E.coli* is facilitated by a distant RNA–DNA hybrid. *EMBO J* 9:2341–2348
- Skoko D, Yoo D, Bai H, Schnurr B, Yan J, McLeod SM, Marko JF, Johnson RC (2006) Mechanism of chromosome compaction and looping by the *Escherichia coli* nucleoid protein Fis. *J Mol Biol* 364:777–798
- Snoep JL, van der Weijden CC, Andersen HW, Westerhoff HV, Jensen PR (2002) DNA supercoiling in *Escherichia coli* is under tight and subtle homeostatic control, involving gene-expression and metabolic regulation of both topoisomerase I and DNA gyrase. *Eur J Biochem* 269:1662–1669
- Stuger R, Woldringh CL, van der Weijden CC, Vischer NO, Bakker BM, van Spanning RJ, Snoep JL, Westerhoff HV (2002) DNA supercoiling by gyrase is linked to nucleoid compaction. *Mol Biol Rep* 29:79–82
- Thompson JF, Landy A (1988) Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. *Nucleic Acids Res* 16:9687–9705

- Thompson JF, Moitoso de Vargas L, Koch C, Kahmann R, Landy A (1987) Cellular factors couple recombination with growth phase: characterization of a new component in the lambda site-specific recombination pathway. *Cell* 50:901–908
- Travers AA, Muskhelishvili G (1998) DNA microloops and microdomains – a general mechanism for transcription activation by torsional transmission. *J Mol Biol* 279:1027–1043
- Travers A, Muskhelishvili G (2005a) DNA supercoiling – a global transcriptional regulator for enterobacterial growth? *Nat Rev Microbiol* 3:157–169
- Travers A, Muskhelishvili G (2005b) Bacterial chromatin. *Curr Opin Genet Dev* 15:507–514
- Travers A, Muskhelishvili G (2007) A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep* 8:147–151
- Travers A, Schneider R, Muskhelishvili G (2001) DNA supercoiling and transcription in *Escherichia coli* – the FIS connection. *Biochimie* 83:213–217
- Tse-Dinh YC, Qi H, Menzel R (1997) DNA supercoiling and bacterial adaptation: thermotolerance and thermoresistance. *Trends Microbiol* 5:323–326
- Ussery D, Larsen TS, Wilkes KT, Friis C, Worning P, Krogh A, Brunak S (2001) Genome organisation and chromatin structure in *Escherichia coli*. *Biochimie* 83:201–212
- Valens M, Penaud S, Rossignol M, Cornet F, Boccard F (2004) Macrodome organization of the *Escherichia coli* chromosome. *EMBO J* 23:4330–4341
- van Drunen CM, van Zuylen CE, Mentjes J, Goosen, van de Putte P (1993) Inhibition of bacteriophage Mu transposition by Mu repressor and Fis. *Mol Microbiol* 10:293–298
- van Workum M, van Dooren SJM, Oldenburg N, Molenaar D, Jensen PR, Snoep JL, Westerhoff HV (1996) DNA supercoiling depends on the phosphorylation potential in *Escherichia coli*. *Mol Microbiol* 20:351–360
- Verbeek H, Nilsson L, Baliko G, Bosch L (1990) Potential binding sites of the trans-activator FIS are present upstream of all rRNA operons and of many but not all tRNA operons. *Biochim Biophys Acta* 1050:302–306
- Verbeek H, Nilsson L, Bosch L (1992) The mechanism of trans-activation of the *Escherichia coli* operon *thrU(tufB)* by the protein FIS. A model. *Nucleic Acids Res* 20:4077–4081
- von Freiesleben U, Rasmussen KV (1992) The level of supercoiling affects the regulation of DNA replication in *Escherichia coli*. *Res Microbiol* 143:655–663
- Wackwitz B, Bongaerts J, Goodman SD, Uuden G (1999) Growth phase-dependent regulation of *nucA-N* expression in *Escherichia coli* K-12 by the Fis protein: upstream binding sites and bioenergetic significance. *Mol Gen Genet* 262:876–883
- Walker KA, Mallik P, Pratt TS, Osuna R (2004) The *Escherichia coli* Fis promoter is regulated by changes in the levels of its transcription initiation nucleotide CTP. *J Biol Chem* 279:50818–50828
- Weinstein-Fischer D, Altuvia S (2007) Differential regulation of *Escherichia coli* topoisomerase I by Fis. *Mol Microbiol* 63:1131–1144
- Weinstein-Fischer D, Elgrably-Weiss M, Altuvia S (2000) *Escherichia coli* response to hydrogen peroxide: a role for DNA supercoiling, topoisomerase I and Fis. *Mol Microbiol* 35:1413–1420
- Wold S, Crooke E, Skarstad K (1996) The *Escherichia coli* Fis protein prevents initiation of DNA replication from *oriC* in vitro. *Nucleic Acids Res* 24:3527–3532
- Wu H, Tyson KL, Cole JA, Busby SJ (1998) Regulation of transcription initiation at the *Escherichia coli* *nir* operon promoter: a new mechanism to account for co-dependence on two transcription factors. *Mol Microbiol* 27:493–505
- Xu J, Johnson RC (1995a) Identification of genes negatively regulated by Fis: Fis and RpoS comodulate growth-phase-dependent gene expression in *Escherichia coli*. *J Bacteriol* 177:938–947
- Xu J, Johnson RC (1995b) Fis activates the RpoS-dependent stationary-phase expression of *proP* in *Escherichia coli*. *J Bacteriol* 177:5222–5231
- Xu J, Johnson RC (1997) Activation of RpoS-dependent *proP* P2 transcription by the Fis protein in vitro. *J Mol Biol* 270:346–359

- Yuan HS, Finkel SE, Feng JA, Kacsor-Grzeskowiak M, Johnson RC, Dickerson RE (1991) The molecular structure of wild-type and mutant Fis protein: relationship between mutational changes and recombinational enhancer function or DNA binding. *Proc Natl Acad Sci USA* 88:9558–9562
- Zacharias M, Göringer HU, Wagner R (1992) Analysis of the Fis-dependent and Fis-independent transcription activation mechanisms of the *Escherichia coli* ribosomal RNA promoter. *Biochemistry* 31:2621–2628
- Zechiedrich EL, Khodursky AB, Cozzarelli NR (1997) Topoisomerase IV, not gyrase, decatenates products of site-specific recombination in *Escherichia coli*. *Genes Dev* 11:2580–2592
- Zechiedrich EL, Khodursky AB, Bachellier S, Schneider R, Chen D, Lilley DMJ, Cozzarelli NR (2000) Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli*. *J Biol Chem* 275:8103–8113

Chapter 15

LRP: A Nucleoid-Associated Protein with Gene Regulatory Functions

Stacey N. Peterson and Norbert O. Reich

Abstract The *Escherichia coli* Leucine-responsive regulatory protein (Lrp) is a nucleoid-associated protein involved in the transcriptional regulation of 10% of all *E. coli* genes. The structural and biochemical properties of Lrp suggest roles in both the organization of the bacterial chromosome as well as transcriptional regulation. Lrp adopts higher order oligomerization states and shows low sequence specificity, similar to other nucleoid-associated proteins. However, Lrp transcriptionally regulates several *E. coli* genes, including those involved in metabolic pathways and those involved in virulence. As its name suggests, leucine affects the ability of Lrp to bind certain sequences and regulate certain genes yet many Lrp-regulated genes are not affected by the presence of leucine.

Lrp homologs are found in both Gram-positive and Gram-negative bacteria and among archaea. Crystal structures of *E. coli* Lrp and various homologs have revealed an octameric structure with conserved motifs in the N and C termini. The role of Lrp in transcription is influenced by such factors as nutrient availability, leucine, interactions with other proteins, and competition with other regulatory proteins for binding specific sequences. Various cellular conditions like nutrient depletion regulate the concentrations of Lrp in the cell which affects the oligomeric state of the protein, influencing its ability to bind to certain sequences.

Keywords Leucine • Lrp • nucleoid-associated protein • transcriptional regulatory protein

S.N. Peterson

Department of Biological Sciences, Mount Saint Mary's College, Los Angeles, California, 90049

N.O. Reich (✉)

Program in Biomolecular Science and Engineering, University of California, Santa Barbara, California, 93106; Department of Chemistry and Biochemistry, University of California, Santa Barbara, California, 93106

e-mail: reich@chem.ucsb.edu

15.1 Functions of Lrp

15.1.1 Role in Transcriptional Regulation

Lrp is often referred to as a global transcriptional regulator because of its role in regulating so many genes. *E. coli* Lrp functions as both an activator and a repressor and is involved in the regulation of more than 400 genes, ~10% of the entire *E. coli* genome (Tani et al. 2002). Most of these genes were identified by microarray analyses and Lrp's role in their regulation is yet to be characterized (Tani et al. 2002). Included in these 400 genes are ~70% of the 215 *E. coli* genes that are induced during the transition from exponential to stationary phase (Tani et al. 2002). Operons under Lrp regulation include those that are involved in amino acid biosynthesis or degradation, nutrient transport, and formation of pili (Calvo and Matthews 1994; Newman and Lin 1995). Lrp directly regulates transcription of many of these genes by binding to sequences within or near their promoters (Calvo and Matthews 1994; Newman and Lin 1995).

In the well-characterized *pap* operon, Lrp acts as both a transcriptional activator and a repressor depending on where it is bound (Hernday et al. 2002). The *pap* operon codes for pyelonephritis-associated pili that allow bacteria to bind to epithelial cells in the urinary tract leading to infection (Johnson 1991). There are six Lrp binding sites in the *pap* regulatory region each containing the consensus GN₂₋₃TTT (Nou et al. 1995). Two binding sites contain GATC sites that may be methylated by DNA adenine methyltransferase (Dam) (Blyn et al. 1990, Figure 15.1a). The GATC sites are often referred to as GATC^{prox} for proximal to *papBA* genes and GATC^{dist} for distal to *papBA* genes. Lrp binds to either sites 1–3 or sites 4–6 (often referred to as half-sites) depending on the methylation state of each GATC and the presence of PapI (Hernday et al. 2003). Methylation of a GATC-containing Lrp binding site decreases the affinity of Lrp for binding that particular half site and favors Lrp binding the other half site (Hernday et al. 2003). The *papBA* genes (those that include the pilus assembly and structural proteins) are actively transcribed when Lrp is bound to sites 4–6 and GATC^{prox} is methylated (Hernday et al. 2002; Fig. 15.1b). Likewise, the *papBA* genes are transcriptionally repressed when Lrp is bound to sites 1–3 and GATC^{dist} is methylated (Hernday et al. 2002). When Lrp is bound to sites 1–3, RNA polymerase cannot gain access to the -10 and -35 promoter sequences and the *papBA* genes are not transcribed. Transcription of the *papBA* genes therefore depends on a competition between Dam and Lrp for access to their binding sites in the *pap* regulatory region. Methylation of one of the two GATC sites results in Lrp assembling at the other site, leading to either activation or repression.

Other proteins are involved in regulating the *pap* operon and some of these proteins interact with Lrp. Leucine, however, is not a regulator of Lrp function in the *pap* operon (Braaten et al. 1992). The local regulator of the *pap* operon, PapI, binds to Lrp and DNA, favouring Lrp binding to sites 4–6 and activation of the pilus genes (Hernday et al. 2003; Fig. 15.1b). The *pap* regulatory region contains two binding sites for PapB, another local regulator (Hernday et al. 2002, Fig. 15.1a). PapB binding

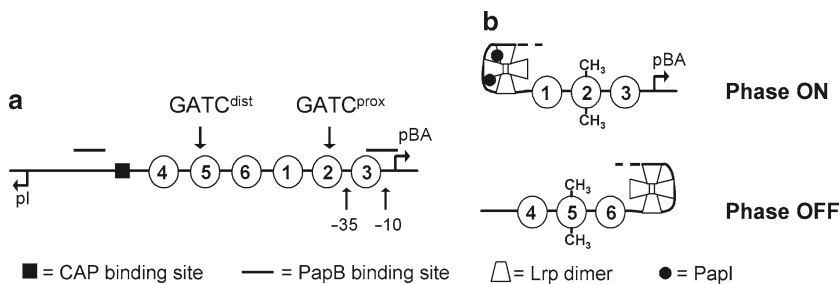


Fig. 15.1 (a) *pap* regulatory region with Lrp binding sites as numbered circles (Hernday et al. 2002). Note: image is not drawn to scale. Lrp binding sites 5 and 2 overlap with GATC sites that may be methylated by Dam. The *pap* regulatory region contains two promoters, pI and pBA. The RNA polymerase -35 and -10 binding sites are indicated for pBA. The CAP binding site and two PapB binding sites are indicated. (b) When Lrp binds sites 4–6 cooperatively with the help of PapI and GATC^{prox} is methylated by Dam, the *papBA* genes are expressed and the bacteria produce pili (phase ON) (Hernday et al. 2002). When Lrp binds sites 1–3 and GATC^{dist} is methylated by Dam, the *papBA* genes are not expressed and no pili are produced (phase OFF) (Hernday et al. 2002). Lrp binds each individual site as a dimer (Cui et al. 1996) and is predicted to bind each half-site (1–3 or 4–6) as an octamer with one dimer unbound, due to cooperative binding

to its upstream site activates the *papI* promoter, increasing the levels of PapI (Hernday et al. 2002). PapB binding to its downstream site may act to repress transcription of the *papBA* genes (Forsman et al. 1989). No interaction between PapB and Lrp has been shown. CpxR-P, the phosphorylated response regulator of the two-component CpxAR envelope stress system, acts to inhibit *papBA* expression and directly competes with Lrp for binding the *pap* regulatory region (Hernday et al. 2004). The global activator Catabolite Activator Protein (CAP) is required for activation of the *papBA* genes. The CAP binding site lies upstream of all six Lrp binding sites (Fig. 15.1a). A direct CAP-RNA polymerase interaction has been implicated in the activation of the pilus genes (Weyand et al. 2001). Because the CAP binding site is ~ 215 bp upstream from the *papBA* promoter it is likely that this interaction is accomplished through Lrp-induced bending of the DNA (Weyand et al. 2001).

As described above, Lrp is not the sole activator of the *pap* operon but collaborates with other proteins like CAP and PapI. Likewise, its role as an activator depends on its ability to compete with proteins for binding DNA, like Dam and CpxR-P. Interaction and competition with proteins affects the function of Lrp in regulating other operons as well (Calvo and Matthews 1994).

Another well-characterized operon in which Lrp plays a significant role and in which leucine acts as a negative regulator is *ilvIH* (Wang and Calvo 1993a). This operon codes for acetohydroxy acid synthase III which is involved in the biosynthesis of branched-chain amino acids (de Felice et al. 1982). The *ilvIH* promoter resides just upstream of the *ilvIH* genes and contains two sets of Lrp binding sites, an upstream region with two Lrp binding sites and a downstream region of four Lrp binding sites (Wang and Calvo 1993a). One of the upstream Lrp binding sites is

quite similar in sequence to the SELEX-selected Lrp consensus sequence (Cui et al. 1995). Cooperative Lrp binding to both regions increases the expression of the *ilvIH* genes (Jafri et al. 2002). Therefore, Lrp acts as a positive regulator of the *ilvIH* operon. Leucine is a negative regulator of the *ilvIH* operon and lowers the affinity of Lrp for its sites in the promoter (Ricca et al. 1989). Under certain conditions the nucleoid-associated protein H-NS prevents Lrp-induced activation of the *ilvIH* genes as well (Levinthal et al. 1994). H-NS has also been shown to have a synergistic relationship with Lrp in regulating rRNA transcription in *E. coli* (Pul et al. 2005).

The fact that only a subset of Lrp-regulated operons is influenced by leucine is particularly intriguing. Leucine favours dissociation of the Lrp hexadecamer to a leucine-bound octamer (Chen and Calvo 2002) suggesting that some promoters require Lrp to adopt a particular oligomerization state (octamer or hexadecamer) in order to bind, while for others either oligomeric state is sufficient. For example, Lrp may bind as a hexadecamer or leucine-bound octamer to some promoters without a preference for one form while for other promoters Lrp may only bind and function as a regulator if it is a hexadecamer *or* a leucine-bound octamer (Chen and Calvo 2002). This hypothesis would suggest that Lrp binds the *ilvIH* promoter as a hexadecamer and the addition of leucine, favouring the octamer, would result in a lower binding affinity and the inability of Lrp to activate. This is not the case, however, and a mutant Lrp that exists as an octamer and cannot assemble into a hexadecamer is still able to activate the *ilvIH* genes (Chen et al. 2005). It has been suggested that the leucine-bound octamer may act as a competitive inhibitor of the non-leucine bound octamer for binding DNA and only the non-leucine bound octamer can activate transcription of the *ilvIH* genes (Chen et al. 2005). Therefore, perhaps both leucine-bound or unbound Lrp octamer bind some regulatory sequences with a similar affinity and regulate identically while for others the leucine-bound octamer may bind less efficiently or perhaps bind well but not be able to function as a regulator. As discussed in the next section, Lrp binding sites within different promoters are degenerate suggesting that the individual Lrp binding sequences and the spacing between the sites may play a large role in whether leucine influences Lrp function.

15.1.2 Role in DNA Organization

Besides its role in transcriptional regulation, Lrp also plays a role in organizing the nucleoid structure of the bacterial chromosome (D'Ari et al. 1993; Luijsterburg et al. 2006). Various Lrp homologs when bound to DNA cause dramatic DNA conformational changes including looping, bridging of DNA molecules, and topology changes (Jafri et al. 1999; Tapias et al. 2000; Beloin et al. 2003). Both atomic force microscopy (AFM) and electron microscopy have revealed structures of multimeric Lrp homologs with DNA wrapped around them, similar to eukaryotic DNA wrapped around a histone core (Beloin et al. 2003). The crystal structures of *E. coli*

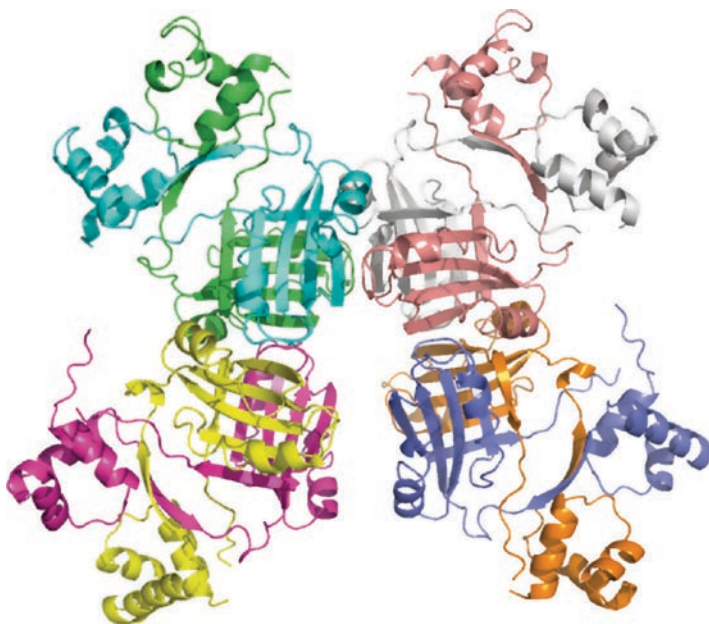


Fig. 15.2 Crystal structure of the *E. coli* Lrp octamer (courtesy of Dr. John Perona, University of California, Santa Barbara) with each monomer a different color (de los Rios and Perona 2007, pdb accession number 2GQQ)

Lrp and its homologs reveal octameric structures with the DNA binding domains at the surface, further supporting the role of Lrp in packaging DNA (Leonard et al. 2001; Thaw et al. 2006; de los Rios and Perona 2007) (Fig. 15.2).

Footprinting studies used to determine Lrp binding sites in various promoters have often revealed large nucleoprotein complexes (sometimes extending over 100 basepairs) that are protected from cleavage (Gazeau et al. 1994; Wiese et al. 1997; McFarland et al. 2008). For example, Lrp binding to the six sites within the *ilvIH* promoter results in the formation of a large nucleoprotein complex not available for cleavage by DNase (Wang and Calvo 1993b). Circular permutation studies revealed that Lrp induces a DNA bend of $\sim 55^\circ$ when bound to site 2 in the *ilvIH* promoter, and induces a bend of $\sim 135^\circ$ when bound to sites 1 and 2 in the promoter (Wang and Calvo 1993b).

Compared to most regulatory proteins, Lrp exists in the cell at a relatively high concentration of $\sim 6,000$ molecules per cell corresponding to $\sim 15 \mu\text{M}$ monomer concentration (Willins et al. 1991). This in vivo concentration is however, much less than typical nucleoid-associated proteins such as H-NS and Fis, which can exist at up to $\sim 20,000$ and $\sim 60,000$ molecules per cell, respectively (Azam et al. 1999). The concentration of Lrp fluctuates depending on the cellular environment and growth phase. For example, the *lrp* gene is upregulated in the presence of guanosine tetraphosphate (ppGpp) (Landgraf et al. 1996). Lrp concentrations in vivo are lowest during late exponential phase when ppGpp levels are low.

Although Lrp is not as abundant as the typical nucleoid-associated proteins, its dramatic effects on DNA structure suggest a role in DNA organization. Furthermore, the concentrations of the typical nucleoid-associated proteins H-NS and Fis fluctuate as well, with Fis even dropping to levels close to zero during late exponential phase (Azam et al. 1999).

Another characteristic of a nucleoid-associated protein is the ability to bind DNA in a sequence-independent manner. Lrp shows a low discrimination between binding specific and nonspecific DNA and its cooperative binding is not significantly affected when bound to nonspecific DNA (Peterson et al. 2007). Various Lrp consensus sequences have been identified. The 15 basepair SELEX-derived consensus sequence (C/T)AG(A/C/T)A(A/T)ATT(A/T)T(A/G/T)CT(A/G) is very degenerate but has helped in the identification of Lrp binding sites within bacterial promoters (Cui et al. 1995). In the *pap* operon the consensus sequence for the six Lrp binding sites in the regulatory region is GN₂₋₃TTT (Nou et al. 1995). The SELEX method selected for short specific sequences that were favourably bound by one Lrp dimer. Therefore, cooperative binding of several Lrp dimers assembling on DNA was not incorporated in the identification of the consensus sequence (Cui et al. 1995). Because Lrp binds multiple sites within regulatory regions by forming higher order structures, the identification of a universal consensus sequence remains complicated. Many promoters that Lrp has been shown to bind and regulate do not contain the SELEX-selected consensus sequence; for some of these the minimal DNA sequence contacted by Lrp has not yet been determined. The Lrp homologue LrpC from *Bacillus subtilis* has a greater binding affinity for a curved region 5' of the *lrpC* gene that did not contain the SELEX-selected consensus compared to a region where the consensus was present (Beloin et al. 2000). This study and others have demonstrated that *E. coli* Lrp and certain homologues prefer binding to regions that are intrinsically curved, consisting of multiple A/T basepairs (Beloin et al. 2000; Tapias et al. 2000; Peterson et al. 2007). The affinity for nonspecific DNA decreases with increasing G/C content (Peterson et al. 2007). It appears that the binding of Lrp to nonspecific DNA requires a minimal length of DNA, perhaps enough so each dimer of an octamer is provided with a binding site. For example, Lrp did not bind short (~64 bp) nonspecific sequences (Azam and Ishihama 1999) but bound nonspecific DNA of larger lengths (~250 bp) (Peterson et al. 2007).

15.2 Structure of Lrp

An Lrp monomer is roughly 18.8 kDa and includes an N terminal DNA binding domain and a C terminal domain responsible for leucine binding and oligomerization interactions (Willins et al. 1991; Platko and Calvo 1993; Chen et al. 2001b). The N-terminal DNA binding domain contains a helix-turn-helix motif (Willins et al. 1991). These motifs are common in transcriptional regulatory proteins that bind in a sequence-specific manner within the major groove of DNA. The C terminus contains both α helices and β sheets, and includes a Regulation of Amino acid

Metabolism (RAM) domain where leucine binds (Ettema et al. 2002; de los Rios and Perona 2007).

Four Lrp homolog crystal structures have been solved which include the *E. coli* AsnC, the *Bacillus subtilis* LrpC, and two homologs in archaea (LrpA and FL11) (Leonard et al. 2001; Koike et al. 2004; Thaw et al. 2006). The AsnC structure was solved in the presence of its effector, asparagine (Thaw et al. 2006). Most recently the *E. coli* Lrp structure has been solved in the presence of DNA although the DNA was not observed in the structure (de los Rios and Perona 2007). The *E. coli* Lrp and all homolog structures except for FL11 reveal an octameric structure with the N terminal DNA binding domains at the surface and the C terminal domains internal, allowing for interactions between each monomer. The *E. coli* Lrp structure shows an open octamer in contrast to the closed homolog structures, perhaps reflecting the influence of the DNA in the crystal (de los Rios and Perona 2007; Figure 15.2).

The Lrp hexadecamer which was shown to be the dominant oligomerization state at high micromolar concentrations in solution (Chen et al. 2001b) has not been revealed by X-ray crystallography. Likewise, the tetramer of LrpC (an Lrp homolog) which by gel filtration studies was shown to be the predominant oligomeric state at low micromolar concentrations has not been revealed by crystallography (Tapias et al. 2000; Thaw et al. 2006). This suggests that although these oligomerization states may be prevalent under certain conditions in solution they may not form stable crystals and are more difficult to crystallize.

15.2.1 Oligomeric States and Influence by the Environment

E. coli Lrp exists in solution as a dimer, octamer or hexadecamer (Chen et al. 2001b). The predominant form depends on the concentration of Lrp and the cellular environment, with dimers predominant at nanomolar concentrations and octamers and hexadecamers at micromolar Lrp concentrations (Chen et al. 2001b). The octameric or hexadecamer states are likely to be the functional units due to the arrangement of Lrp binding sites in many regulatory regions and the evidence that Lrp binds an individual site as a dimer (Cui et al. 1996). Furthermore, Lrp concentrations in vivo are in the micromolar range (~15 μ M) so octamers and hexadecamers would be most prevalent (Willins et al. 1991). Gel mobility shifts performed with low nanomolar Lrp concentrations show cooperative binding of Lrp on DNA suggesting that the presence of DNA may help increase the formation of large oligomerization states when Lrp concentrations are low (Nou et al. 1995; Hernday et al. 2003; Chen et al. 2005).

The concentration of Lrp in the cell is higher in minimal media during mid-log phase compared to rich media (Chen et al. 2001a). In minimal media, Lrp concentration does not fluctuate much between log and stationary phase (Landgraf et al. 1996). In rich media however, more Lrp is present during stationary phase than mid-log (Landgraf et al. 1996). Environmental conditions also affect the amount of

Lrp that is bound to DNA *in vivo*; in minimal media, less Lrp is bound to DNA than in rich media (Chen et al. 2001a). *In vivo* studies using minicells revealed nonspecific dissociation constants in the millimolar range, much higher than the nanomolar values calculated from *in vitro* studies (Chen et al. 2001a; Peterson et al. 2007). This difference probably results from the difference in nucleoid DNA versus naked DNA *in vitro*, where *in vivo* Lrp has a lot more competition for binding DNA due to other nucleoid-associated and regulatory proteins. As mentioned earlier, Lrp regulates several genes (~10% of total *E. coli* genes) with its role in regulation not characterized for many of these genes. However, for the metabolic operons regulated by Lrp, Lrp generally activates transcription of genes involved in anabolic pathways and represses those involved in catabolic pathways (Calvo and Matthews 1994). This correlates with its condition-dependent variation in percent bound and free; in minimal media where catabolic pathways predominate less Lrp is bound to DNA, and in rich media where anabolic pathways predominate more Lrp is bound to DNA (Chen et al. 2001a).

It is proposed that growth conditions (minimal versus rich media) affect the oligomerization states of Lrp, which in turn affects the affinity of Lrp for various DNA sequences (Chen et al. 2001a). Rich media containing leucine would favour the leucine-bound octamer at *in vivo* concentrations. The leucine-bound octamer may have a higher affinity for nonspecific DNA as suggested from minicell experiments (Chen et al. 2001a) which may further decrease its binding affinity for certain regulatory DNA elements, such as those in the *ilvIH* regulatory region. Further complicating the interpretations is the finding through dynamic light-scattering and fluorescent studies that Lrp contains both a high-affinity and a low-affinity leucine binding site, both residing within the C terminal region of the protein (Chen and Calvo 2002). Leucine binding to the low affinity site induces Lrp hexadecamers to dissociate to octamers (Chen et al. 2001b; Chen and Calvo 2002). The effect of leucine binding the high affinity site has not been characterized but likely causes a conformational change in Lrp; this change may contribute to its binding specificity for certain sites (Chen and Calvo 2002).

15.3 Lrp Homologues

Lrp homologs are found in eubacteria and archaea and Lrp is highly conserved within enteric bacteria (Friedberg et al. 1995). *In vivo* Lrp concentrations of some of these homologs vary dramatically from *E. coli* Lrp. For example, the Lrp homolog in *H. influenzae* is present *in vivo* at only ~130 dimers (compared to in *E. coli* where 3,000 dimers are present) (Friedberg et al. 2001). This finding has led to the conclusion that this particular homolog does not play a role in DNA organization due to its limited concentration *in vivo* and is involved in regulating a few select genes as opposed to several diverse genes (Friedberg et al. 2001).

Several homologs of Lrp have been shown to be functionally dependent on amino acids other than leucine, including AsnC which is regulated by asparagine

levels (de Wind et al. 1985; Kolling and Lother 1985). AsnC, which shares 25% identity with Lrp (Willins et al. 1991), regulates its own expression and that of *asnA*, which codes for asparagine synthetase (Kolling and Lother 1985). High asparagine concentrations prevent AsnC from activating transcription of *asnA*, decreasing the levels of asparagine synthetase (Kolling and Lother 1985). PutR, an Lrp homologue with 47% identity to *E. coli* Lrp (Cho and Winans 1996), is a transcriptional activator of the *putA* gene which is required for catabolism of proline in *Agrobacterium tumefaciens*. Proline influences the binding of PutR and its ability to activate expression of *putA* (Jafri et al. 1999). Like Lrp, PutR has been shown through AFM and circular permutation studies to bend DNA although its binding is not largely cooperative (Jafri et al. 1999). Other Lrp homologues that are bound and regulated by amino acids include the FL11 from *Pyrococcus* which binds glutamine (Koike et al. 2004), BkdR from *Pseudomonas putida* which binds valine (Madhusudhan et al. 1997), and LysM from *Sulfolobus solfataricus*, which binds lysine (Brinkman et al. 2002).

LrpC from *Bacillus subtilis* shares 34% identity with *E. coli* Lrp (Beloin et al. 1997). Unlike *E. coli* Lrp, the predominant form of LrpC in solution at 6 μ M is a tetramer (Tapias et al. 2000) as opposed to an octamer or hexadecamer (Chen et al. 2001b). Although LrpC is present at much lower concentrations in the cell than *E. coli* Lrp (50–300 molecules per cell) (Beloin et al. 2000) which argues against a role in DNA organization, LrpC promotes DNA bending and increases the binding affinity of the *B. subtilis* HU-like nucleoid-associated protein Hbsu for DNA (Tapias et al. 2000). Furthermore, both AFM and electron microscopy studies have revealed the wrapping of DNA around LrpC similar to eukaryotic nucleosomes (Beloin et al. 2003). These properties support a role of LrpC in nucleoid organization.

15.4 Future Research on Lrp

Many Lrp regulated genes are induced during the transition from log phase to stationary phase where Lrp concentrations peak (Tani et al. 2002). Most of these genes have yet to be characterized. It is important to understand how Lrp contributes to the regulation of such genes. For example, does Lrp bind to the regulatory region and affect the formation of the transcription initiation complex or does Lrp act indirectly through activating or repressing another protein that is needed for regulating the gene? Furthermore, many Lrp-regulated operons remain poorly understood.

An essential question concerning Lrp and its role in both nucleoid organization and gene regulation is how the distinct Lrp oligomerization states impact the protein's function. This has implications far beyond Lrp since many proteins which assemble onto DNA undergo such transitions. For Lrp, this is particularly important since small molecules (e.g., leucine), proteins (e.g., PapI), and DNA context (e.g., number of Lrp consensus sites, nature of available sequences) profoundly influence this assembly

process. A related and largely structural question is how does Lrp assemble onto particular DNA sequences in the context of other proteins which also bind that particular locus? Few examples (*pap*, *ilvIH*) are fairly well understood. This is a question which is particularly challenging, not only for those interested in Lrp or even bacterial nucleoid-associated proteins, because it falls into the experimental twilight zone between classical methods (e.g., X-ray crystallography and NMR) and less direct methods (e.g., FRET – Fluorescence Resonance Energy Transfer). Of particular relevance are those methods amenable to the study of complexes involving DNA segments in the 100–150 bp range, such as AFM. Finally, an understanding of such static complexes, although enormously helpful when they do finally emerge, will still leave unanswered the issues of how such assemblies actually form. This requires a functional dissection relying on kinetic measurements coupled with spatial information. To date, the best method for such studies would seem to be single molecule FRET, which has yet to be applied to such challenging studies.

Another interesting area of future Lrp research is to understand better the evolutionary role of Lrp and which function (regulation or DNA organization) originated first. Was Lrp a nucleoid-associated protein that evolved into a regulatory protein or vice versa? This question may be approached by bioinformatics analyses looking at homology between proteins as well as structural approaches (more crystal structures of homologs) and sequence and structural comparisons with other established nucleoid-associated proteins.

References

- Azam TA, Ishihama A (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* 274: 33105–33113
- Azam TA, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181:6361–6370
- Beloin C, Ayora S, Exley R, Hirschbein L, Ogasawara N, Kasahara Y, Alonso JC, Le Hegarat F (1997) Characterization of an *lrp*-like (*lrpC*) gene from *Bacillus subtilis*. *Mol Gen Genet* 256:63–71
- Beloin C, Exley R, Mahe A, Zouine M, Cubasch S, Le Hegarat F (2000) Characterization of LrpC DNA-binding properties and regulation of *Bacillus subtilis lrpC* gene expression. *J Bacteriol* 182:4414–4424
- Beloin C, Jeusset J, Revet B, Mirambeau G, Le Hegarat F, Le Cam E (2003) Contribution of DNA conformation and topology in right-handed DNA wrapping by the *Bacillus subtilis* LrpC protein. *J Biol Chem* 278:5333–5342
- Blyn LB, Braaten BA, Low DA (1990) Regulation of *pap* pilin phase variation by a mechanism involving differential dam methylation states. *EMBO J* 9:4045–4054
- Braaten BA, Platko JV, van der Woude MW, Simons BH, De Graaf FK, Calvo JM, Low DA (1992) Leucine-responsive regulatory protein controls the expression of both the *pap* and *fan* pili operons in *Escherichia coli*. *Proc Natl Acad Sci USA* 89:4250–4254
- Brinkman AB, Bell SD, Lebbink RJ, de Vos WM, van der Oost J (2002) The *Sulfolobus solfataricus* Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability. *J Biol Chem* 277:29537–29549

- Calvo JM, Matthews RG (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* 58:466–490
- Chen S, Calvo JM (2002) Leucine-induced dissociation of *Escherichia coli* Lrp hexadecamers to octamers. *J Mol Biol* 318:1031–1042
- Chen S, Hao Z, Bieniek E, Calvo JM (2001a) Modulation of Lrp action in *Escherichia coli* by leucine: effects on non-specific binding of Lrp to DNA. *J Mol Biol* 314:1067–1075
- Chen S, Rosner MH, Calvo JM (2001b) Leucine-regulated self-association of leucine-responsive regulatory protein (Lrp) from *Escherichia coli*. *J Mol Biol* 312:625–635
- Chen S, Iannolo M, Calvo JM (2005) Cooperative binding of the leucine-responsive regulatory protein (Lrp) to DNA. *J Mol Biol* 345:251–264
- Cho K, Winans SC (1996) The *putA* gene of *Agrobacterium tumefaciens* is transcriptionally activated in response to proline by an Lrp-like protein and is not autoregulated. *Mol Microbiol* 22:1025–1033
- Cui Y, Wang Q, Stormo GD, Calvo JM (1995) A consensus sequence for binding of Lrp to DNA. *J Bacteriol* 177:4872–4880
- Cui Y, Midkiff MA, Wang Q, Calvo JM (1996) The leucine-responsive regulatory protein (Lrp) from *Escherichia coli*. Stoichiometry and minimal requirements for binding to DNA. *J Biol Chem* 271:6611–6617
- D'Ari R, Lin RT, Newman EB (1993) The leucine-responsive regulatory protein: more than a regulator? *Trends Biochem Sci* 18:260–263
- De Felice M, Lago CT, Squires CH, Calvo JM (1982) Acetohydroxy acid synthase isoenzymes of *Escherichia coli* K12 and *Salmonella typhimurium*. *Ann Microbiol (Paris)* 133:251–256
- de los Rios S, Perona JJ (2007) Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *J Mol Biol* 366:1589–1602
- De Wind N, de Jong M, Meijer M, Stuitje AR (1985) Site-directed mutagenesis of the *Escherichia coli* chromosome near *oriC*: identification and characterization of *asnC*, a regulatory element in *E. coli* asparagine metabolism. *Nucleic Acids Res* 13:8797–8811
- Ettema TJ, Brinkman AB, Tani TH, Rafferty JB, van der Oost J (2002) A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes. *J Biol Chem* 277:37464–37468
- Forsman K, Goransson M, Uhlin BE (1989) Autoregulation and multiple DNA interactions by a transcriptional regulatory protein in *E. coli* pili biogenesis. *EMBO J* 8:1271–1277
- Friedberg D, Platko JV, Tyler B, Calvo JM (1995) The amino acid sequence of Lrp is highly conserved in four enteric microorganisms. *J Bacteriol* 177:1624–1626
- Friedberg D, Midkiff M, Calvo JM (2001) Global versus local regulatory roles for Lrp-related proteins: *Haemophilus influenzae* as a case study. *J Bacteriol* 183:4004–4011
- Gazeau M, Delort F, Fromant M, Dessen P, Blanquet S, Plateau P (1994) Structure-function relationship of the Lrp-binding region upstream of *lysU* in *Escherichia coli*. *J Mol Biol* 241:378–389
- Hernday A, Krabbe M, Braaten B, Low D (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci USA* 99(Suppl 4):16470–16476
- Hernday AD, Braaten BA, Low DA (2003) The mechanism by which DNA adenine methylase and PapI activate the *pap* epigenetic switch. *Mol Cell* 12:947–957
- Hernday AD, Braaten BA, Broitman-Maduro G, Engelberts P, Low DA (2004) Regulation of the *pap* epigenetic switch by CpxAR: phosphorylated CpxR inhibits transition to the phase ON state by competition with Lrp. *Mol Cell* 16:537–547
- Jafri S, Evoy S, Cho K, Craighead HG, Winans SC (1999) A Lrp-type transcriptional regulator from *Agrobacterium tumefaciens* condenses more than 100 nucleotides of DNA into globular nucleoprotein complexes. *J Mol Biol* 288:811–824
- Jafri S, Chen S, Calvo JM (2002) *ilvIH* operon expression in *Escherichia coli* requires Lrp binding to two distinct regions of DNA. *J Bacteriol* 184:5293–5300
- Johnson JR (1991) Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev* 4:80–128

- Koike H, Ishijima SA, Clowney L, Suzuki M (2004) The archaeal feast/famine regulatory protein: potential roles of its assembly forms for regulating transcription. *Proc Natl Acad Sci USA* 101:2840–2845
- Kolling R, Lother H (1985) AsnC: an autogenously regulated activator of asparagine synthetase. A transcription in *Escherichia coli*. *J Bacteriol* 164:310–315
- Landgraf JR, Wu J, Calvo JM (1996) Effects of nutrition and growth rate on Lrp levels in *Escherichia coli*. *J Bacteriol* 178:6930–6936
- Leonard PM, Smits SH, Sedelnikova SE, Brinkman AB, de Vos WM, van der Oost J, Rice DW, Rafferty JB (2001) Crystal structure of the Lrp like transcriptional regulator from the archaeon *Pyrococcus furiosus*. *EMBO J* 20:990–997
- Levinthal M, Lejeune P, Danchin A (1994) The H-NS protein modulates the activation of the *ilvIH* operon of *Escherichia coli* K12 by Lrp, the leucine regulatory protein. *Mol Gen Genet* 242:736–743
- Luijsterburg MS, Noom MC, Wuite GJ, Dame RT (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol* 156:262–272
- Madhusudhan KT, Huang N, Braswell EH, Sokatch JR (1997) Binding of L-branched-chain amino acids causes a conformation change in BkdR. *J Bacteriol* 179:276–279
- McFarland KA, Lucchini S, Hinton JCD, Dorman CJ (2008) The leucine-responsive regulatory protein, Lrp, activates transcription of the *fim* operon in *Salmonella enterica* serovar typhimurium via the *fimZ* regulatory gene. *J Bacteriol* 190:602–612
- Newman EB, Lin R (1995) Leucine-responsive regulatory protein: a global regulator of gene expression in *E. coli*. *Annu Rev Microbiol* 49:747–775
- Nou X, Braaten B, Kaltenbach L, Low DA (1995) Differential binding of Lrp to two sets of *pap* DNA binding sites mediated by PapI regulates *Pap* phase variation in *Escherichia coli*. *EMBO J* 14:5785–5797
- Peterson SN, Dahlquist FD, Reich NO (2007) The role of high affinity non-specific DNA binding by Lrp in transcriptional regulation and DNA organization. *J Mol Biol* 369:1307–1317
- Platko JV, Calvo JM (1993) Mutations affecting the ability of *Escherichia coli* Lrp to bind DNA, activate transcription, or respond to leucine. *J Bacteriol* 175:1110–1117
- Pul U, Wurm R, Lux B, Meltzer M, Menzel A, Wagner R (2005) Lrp and H-NS – cooperative partners for transcription regulation at *Escherichia coli* rRNA promoters. *Mol Microbiol* 58:864–876
- Ricca E, Aker D, Calvo JM (1989) A protein that binds to the regulatory region of the *Escherichia coli* *ilvIH* operon. *J Bacteriol* 171:1658–1664
- Tani TH, Khodursky A, Blumenthal RM, Brown PO, Matthews RG (2002) Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci USA* 99:13471–13476
- Tapias A, Lopez G, Ayora S (2000) *Bacillus subtilis* LrpC is a sequence-independent DNA-binding and DNA-bending protein which bridges DNA. *Nucleic Acids Res* 28:552–559
- Thaw P, Sedelnikova SE, Muranova T, Wiese S, Ayora S, Alonso JC, Brinkman AB, Akerboom J, van der Oost J, Rafferty JB (2006) Structural insight into gene transcriptional regulation and effector binding by the Lrp/AsnC family. *Nucleic Acids Res* 34:1439–1449
- Wang Q, Calvo JM (1993a) Lrp, a global regulatory protein of *Escherichia coli*, binds co-operatively to multiple sites and activates transcription of *ilvIH*. *J Mol Biol* 229:306–318
- Wang Q, Calvo JM (1993b) Lrp, a major regulatory protein in *Escherichia coli*, bends DNA and can organize the assembly of a higher order nucleoprotein structure. *EMBO J* 12:2495–2501
- Weyand NJ, Braaten BA, van der Woude M, Tucker J, Low DA (2001) The essential role of the promoter-proximal subunit of CAP in *pap* phase variation: Lrp- and helical phase-dependent activation of *papBA* transcription by CAP from –215. *Mol Microbiol* 39:1504–1522
- Wiese DE, Ernsting BR, Blumenthal RM, Matthews RG (1997) A nucleoprotein activation complex between the leucine-responsive regulatory protein and DNA upstream of the *ghdBDF* operon in *Escherichia coli*. *J Mol Biol* 270:152–168
- Willins DA, Ryan CW, Platko JV, Calvo JM (1991) Characterization of Lrp, and *Escherichia coli* regulatory protein that mediates a global response to leucine. *J Biol Chem* 266:10768–10774

Chapter 16

Extreme DNA Bending: Molecular Basis of the Regulatory Breadth of IHF

Amalia Muñoz, Marc Valls, and Víctor de Lorenzo

Abstract The Integration host factor (IHF) is a heterodimeric, sequence-specific DNA-binding and DNA-bending protein found in many types of eubacteria. The sole function of IHF is to bring about a sharp curvature in the target DNA (up to $\geq 160^\circ$). Such a drastic change in DNA shape has been evolutionarily recruited for controlling a large number of functions that depend on the architecture of given genomic sites. These include the organization of the bacterial nucleoid and the transcriptional control of distinct promoters. The growing availability of bacterial genomes allows a comparative approach to survey the regulatory breadth of IHF in a wider context. In this Chapter, we use the sequence of the IHF protein of the soil bacterium *Pseudomonas putida* as a starting point to examine in detail the basis of the recognition of DNA sequences by this nucleoid-associated protein, in particular the correlation between sequence conservation and DNA interaction for each of the IHF chains. This is greatly facilitated by comparing the protein sequence and the DNA binding specificity of IHF with those of similar proteins HU and the transcription factor 1 (TF1) from bacteriophage SPO1 of *Bacillus subtilis*. Mapping of the fully conserved amino acids and the protein-specific sites for each chain of the corresponding tridimensional structures finely correlates with those sites involved in DNA interactions and maintaining the protein dimer structure. The sequence conservation profile of the DNA-binding regions of these proteins shows that chain

A. Muñoz and V. de Lorenzo (✉)

Systems Biology Program, Centro Nacional de Biotecnología (CSIC), Campus de Cantoblanco, Madrid, 28049, Spain

Víctor de Lorenzo, Systems Biology Program, Centro Nacional de Biotecnología – CSIC, Campus de Cantoblanco, 28049, Madrid, Spain

e-mail: vdlorenzo@cnb.csic.es

M. Valls

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, 08028, Barcelona, Spain

A. Muñoz

Institute for Reference Materials and Measurement, Joint Research Center of the EU, Geel, Belgium

B of IHF is more closely related to HU/TF1 than to chain A of IHF, suggesting a separate evolutionary origin. Furthermore, some features of the DNA recognition mechanism seem to be exclusive to IHF and cannot be fulfilled by HU or TF1 proteins. HU and TF1 can be *embraced* by DNA as IHF can by the action of residues conserved in the three proteins (thereby explaining why HU/TF1 and IHF can be partially replaced by each other). In contrast, only the interactions mediated by *tree-determinants* (i.e. those residues that are specific for each chain of IHF) can afford a high DNA recognition specificity. These analyses highlight the importance of DNA binding versus DNA bending specificities for expansion of the regulatory space of such nucleoid-associated proteins.

Keywords DNA bending • transcriptional control • tree-determinants • protein evolution • TF1 factor

16.1 Introduction: What Is IHF and What It Does

Integration host factor (IHF) is a small (~20 kDa) basic heterodimeric protein that binds and bends DNA specifically and belongs to the family of prokaryotic nucleoid-associated proteins, also comprising HU, FIS and H-NS (Schmid 1990). The protein is highly abundant in the cell (Ditto et al. 1994; Valls et al. 2002) and is strongly conserved among prokaryotes. It consists of two homologous subunits, α and β , encoded by two unlinked genes (*ihfA* and *ihfB*; Haluzi et al. 1991; Weisberg et al. 1996). IHF was originally discovered as a host factor required for the integration of bacteriophage λ into the *Escherichia coli* chromosome (reviewed in Nash and Robertson (1981)). However, it has been involved later in a variety of cellular processes representing nearly all the major DNA functions, such as replication, transcription, site-specific recombination, transposition, partitioning, transfer and packaging into phage (dos Santos and Rodrigues 2005; Nash 1996). IHF functionality relies on its ability to sharply bend DNA (Dixit et al. 2005; Kuznetsov et al. 2006), bringing together sequences separated by several helical turns. Its architectural role is supported by the fact that IHF can be replaced to a certain extent by other DNA-bending proteins or by intrinsically bent DNA (Goodman et al. 1992; Parekh and Hatfield 1996; Pérez-Martín et al. 1994).

Unlike other HU-like proteins, IHF binds to DNA with a high sequence specificity: target sites are preferred over random sequences by two to three orders of magnitude (Wang et al. 1995; Yang and Nash 1995). Thus, the protein produces a distinct footprint on target sequences that typically protects 30 to 35 bp and suggests that the binding site is recognized through contacts with the minor groove of DNA (Friedman et al. 1988; Wang et al. 1995; Yang and Nash 1989). Alignment of the known *E. coli* IHF binding sites revealed a meaningfully conserved sequence element that approximates to the asymmetric consensus WATCAANNNTTR (where W is A or T, R is A or G; Goodrich et al. 1990). Many natural IHF binding

sites also include an A/T tract of 4–6 nucleotides, located approximately eight base pairs upstream from the mentioned consensus, that is devoid of conserved sequence patterns. The presence of this A-rich sequence, located one helical turn upstream from the core consensus creates an intrinsically rigid structure with a narrow minor groove that enhances binding (Goodrich et al. 1990; Hales et al. 1996; Yang and Nash 1989). In *E. coli* the validity of the consensus sequence for IHF sites was confirmed in genetic studies where mutations affecting binding were found within the conserved elements (Lee et al. 1991).

Similarly to HU-like DNA binding proteins, IHF subunits contain a helix-turn-helix domain involved in dimerisation and two antiparallel β -sheets that bind to DNA (Rice et al. 1996; Tanaka et al. 1984). The co-crystal structure of *E. coli* IHF bound to the λ phage H9 site shows the DNA bent by $\sim 180^\circ$ around the heterodimer in a virtual U-turn (Rice 1997; Rice et al. 1996). The side-chains of two proline residues at the turns of separate β -sheet arms intercalate between base-pairs from the minor groove side of the duplex and introduce kinks in the DNA helix (Fig. 16.1). Although the base-pairs adjacent to the intercalation site are severely distorted, the double helix is not melted at the binding site. The general

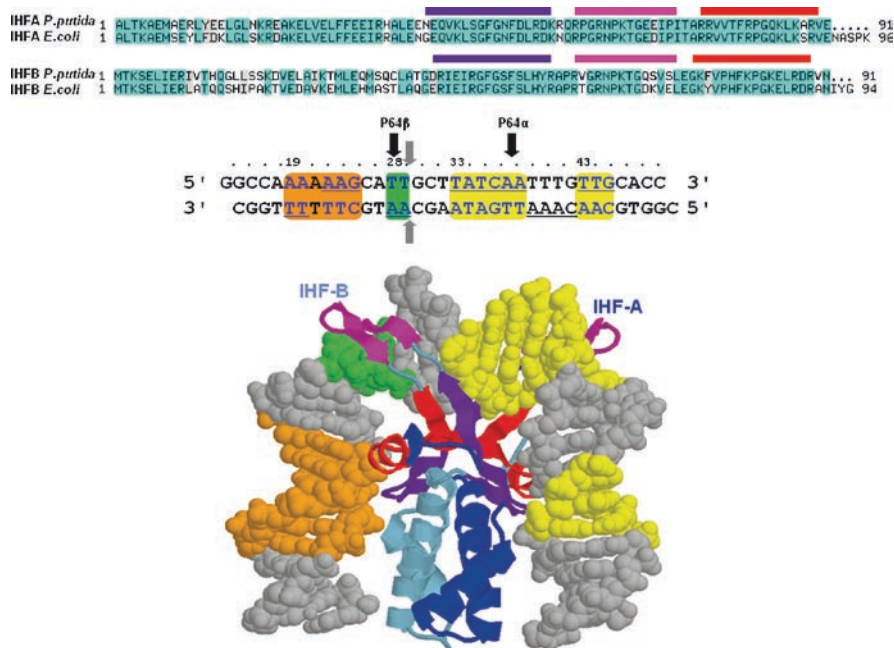


Fig. 16.1 IHF sequences and DNA recognition. *Top*: Alignment showing the degree of sequence homology between IHF from *E. coli* and *P. putida*. The DNA binding HU motifs are indicated in purple, magenta and red. *Bottom*: IHF-protein interactions. The picture shows the main interacting regions within the complex modeled for the *P. putida* protein.

architecture of the DNA-protein interactions does not change whether the DNA in the crystal is an intact synthetic sequence or a nicked DNA. The solved structure indicates that the majority of contacts between IHF and the H9 site are to the phosphates and riboses of the DNA backbone. Only one amino acid Arg-46 in the β -subunit makes a hydrogen bond with a DNA base of the TTR element conserved at the 3' end of the consensus. This amino acid is located in a small β -sheet that is highly conserved among different species and its importance for site recognition had previously been proposed through genetic studies (Yang and Nash 1989). Indeed, mutations in β -Glu44 were isolated in a search for IHF variants that restored binding to a mutant site that contained a T to A change at the central position of the TTR element (Lee et al. 1992). The importance of the TTR element for sequence recognition has been confirmed by comparison of the crystal structures of IHF-DNA complexes containing mutations in the protein or the DNA sequence (Lynch et al. 2003). A more detailed analysis of the way IHF binds DNA specifically with so few direct contacts with the nucleotides of the target DNA sequence is presented below.

16.2 IHF in Transcription Control

IHF has been found to participate in the positive and negative control of gene expression in a number of Gram-negative bacteria (Freundlich et al. 1992). No obvious physiological relationship is found between these genes but transcription of many of them is dependent on the RNA polymerase holoenzyme containing the alternative factor σ^{54} . This holoenzyme always requires the presence of an activator protein that binds to the promoter upstream region in order to promote transcription. The role of IHF is to facilitate the interaction between the enhancer-binding activator and the promoter-bound RNA polymerase- σ^{54} complex to bring about transcriptional activation (Abril and Ramos 1993; de Lorenzo et al. 1991; Pérez-Martín and de Lorenzo 1996a; Wedel et al. 1990). Examples of these systems are the genes for nitrogen fixation in *Klebsiella pneumoniae*, regulation of flagellum synthesis in *Caulobacter crescentus*, alginate production in *Pseudomonas aeruginosa*, and toluene degradation in the *Pseudomonas putida* TOL plasmid pWW0 (de Lorenzo et al. 1991; Delic-Attree et al. 1996; Gober and Shapiro 1990; Hoover et al. 1990). In the latter system, the *Pu* promoter regulating the upper *xyl* operon has been extensively studied both in vivo and in vitro (Abril and Ramos 1993; Calb et al. 1996; Pérez-Martín and de Lorenzo 1996a). IHF binding to *Pu* fixes the optimal promoter geometry, which facilitates contacts between distant proteins and aids the recruitment of σ^{54} -RNA polymerase (Bertoni et al. 1998; Pérez-Martín et al. 1994). Consequently, *Pu* expression in *E. coli* and *P. putida* strains that lack IHF is almost undetectable (Calb et al. 1996; de Lorenzo et al. 1991; Valls et al. 2002). A large number of σ^{70} -dependent promoters are also subject to positive or negative regulation by IHF (Goosen and van de Putte 1995). In some cases (typically the PL promoter of lambda phage) IHF is the only regulator, while in others the factor

works as a co-regulator in combination with other activators or repressors (Pérez-Martín and de Lorenzo 1996b).

Understanding what determines IHF binding and its specific targets in the genome is critical to reveal its physiological role (Senear et al. 2007). Search for IHF sites has been routinely applied within bacteriophage genomes, plasmids, insertion elements and DNA regions of a variety of bacteria (Delic-Attree et al. 1995; Gober and Shapiro 1990; Grainger et al. 2006; Kuznetsov et al. 2006; Lee et al. 1991). This has been often done by looking for a match to degenerate consensus motifs deduced from experimentally defined target sequences. A thorough genome-wide estimate of the binding sites has been performed in the *E. coli* chromosome, where the existence and location of over 600 IHF sites was predicted using an extended consensus and *Hidden Markov Models* (Ussery et al. 2001). The same study also identified 130 additional target sites associated with repetitive elements (Ussery et al. 2001). Nevertheless, the fact that the already known sites -whose sequences have often been identified by similarity to the established consensus are used to train the search might strongly bias the results of these studies.

16.3 IHF is a Nucleoid-Associated Protein

In both eukaryotic and prokaryotic organisms, the disproportion between the size of the cellular compartment hosting the genome and the size of the genome is in part dealt with by organization in loops. These are higher-order structures where the chromatin fibre is folded into topologically independent supercoiled domains, the torsional stress of which is relieved by the presence of nucleosomes (Ussery et al. 2001). Bacterial chromosomes also display a sort of supercoiled loop domain organization, although with a certain degree of torsional stress due to the just partial compensation resulting from the lower amount of proteins involved in genome organization and their less stable interactions with DNA (Ochman and Davalos 2006). There are four well-characterized architectural proteins (nucleoid-associated proteins) in bacterial chromatin HU (Histone-like protein from strain U93), FIS (factor for inversion stimulation), IHF (integration host factor) and H-NS (histone-like nucleoid-structuring protein). These are thought to play an active role in regulatory processes as well as to participate in chromosome compaction as essential structural components (Ali Azam et al. 1999; Ussery et al. 2001). Each of these proteins has specific articles in this book (HU, Chapter 17; FIS, Chapter 14; H-NS, Chapter 13).

From a structural point of view, these four proteins can be differentiated into two different groups depending on their architecture. FIS is composed of four α -helices tightly intertwined to form a globular dimer and it shows an orthogonal bundle structure and a particular topology described as that of FIS proteins. In contrast, HU, IHF and H-NS, exhibit an irregular architecture with few secondary structures and a topology described as that of HU proteins according to CATH (Protein Structural Classification; Pearl et al. 2001). Another differential feature between FIS and the others lies in how they interact with DNA. FIS uses a typical

helix-turn-helix (HTH) surface, while IHF, HU and H-NS employ a histone-like motif. This last *signature pattern* is defined by a 20-residue sequence, which includes three perfectly conserved positions. According to their tertiary structure, this histone-like pattern spans exactly the first half of their flexible DNA-binding arms (see PS00045 motif in <http://www.expasy.ch/prosite>). On the other hand, the HTH motif is a common recognition element in transcriptional regulators, which is typically constituted by 20 amino acids forming two almost perpendicular alpha-helices connected by a loop. This motif invariably binds the DNA major groove, as the second helix, known as the *recognition helix*, is inserted in the groove. In contrast, both HU and IHF type proteins bind the minor groove and disrupt the DNA by intercalating side chains from the beta sheet motifs. HU and IHF act as dimers, a beta-hairpin arm from each subunit extending towards the opposing face of the DNA and inserting proline side chains between distinct base-steps (Nash 1996; Swinger and Rice 2004). The minor groove is thereby widened in the region of binding and the DNA bends toward the main body of the protein. The sections below present a comparative analysis of the modes of binding of IHF and HU (and a third phage protein TF1, see below) as a way to explore the regulatory space of factor-induced DNA bending.

16.4 IHF Versus HU

IHF received its name owing to its ability to facilitate the integration of lambda phage into *E. coli* (Nash and Robertson 1981). The reason for this is now known related *inter alia* to the specific DNA looping of some of the phage promoters (Giladi et al. 1998). By the same token, IHF modulates the transcriptional activity of a large number of promoters by influencing the looping of their upstream DNA. Integration host factor-type proteins are found mostly in enterobacteria and some bacteriophages. HU-type proteins seem to be more widespread in a variety of eubacteria, cyanobacteria and archaeobacteria, as well as in the chloroplast genome of some algae. The physiological functions of these proteins often referred to as histone-like proteins are diverse. Both HU and IHF are not only capable of bending DNA (see Chapter 16), but they also protect it from denaturation under harsh environmental conditions.

Both IHF and HU are closely related proteins of about 20 kDa in size that serve as multipurpose benders of DNA, thereby playing a global role in chromosomal organization. The functional IHF protein exists as a heterodimer of homologous (but not identical) subunits, where each monomer has a distinct role in its interaction with DNA (see below). In contrast, HU proteins work mostly as homodimers, a feature that they share with the related transcription factor 1 (TF1) from bacteriophage SPO1 of *Bacillus subtilis*. Yet there are cases (e.g. *E. coli*), where coexistence of two HU variants in the same bacterium yields functional heterodimers. In fact, the HU protein of *E. coli* seems to work equally well as heterodimer and as homodimer (Claret and Rouvière-Yaniv 1997). The availability of crystal structures

for archetypal proteins of this sort (reviewed in Swinger and Rice, 2004) provide revealing similarities and differences in the mode of binding of these factors to DNA. In all cases (IHF, HU, TF1) the two beta-arms of their structures function as binding surfaces for bacterial DNA. However, although HU, TF1 and IHF share similar DNA binding folds, the HU protein shows little sequence preference. Furthermore, DNA does not fold permanently around the protein, although a transient architecture of this sort can be captured in the crystal structure (Koh et al. 2008). Besides these features, which will be discussed in detail later in this Chapter, IHF and HU have specific roles in genetic recombination and transcriptional control of distinct promoters. Since the loss of such central functions is systematically deleterious for cells, these nucleoid-associated proteins are essential for the survival of pathogenic and commensal microbes inside the human host. Furthermore, in some bacterial species, IHF and/or HU genes contribute directly to virulence (Stonehouse et al. 2008). The intracellular concentration of these proteins change during bacterial cell growth, IHF increasing during stationary phase (Valls et al. 2002) and HU changing both its intracellular concentration and its subunit composition (Claret and Rouvi  re-Yaniv 1997). Finally, these proteins seem to assist the compacting of the chromosome during stationary growth phase (Ali et al. 2001; Ussery et al. 2001). It thus comes as no surprise that IHF and HU mutants present very pleiotropic phenotypes (Painbeni et al. 1997).

16.5 How IHF Binds and Bends Its Target DNA

Numerous studies have suggested that the mechanism of protein-DNA recognition relies not just on direct recognition of base-pairs, but also on indirect interactions, e.g. sequence-dependent structural features of the DNA, such as backbone conformation and flexibility (Aeling et al. 2007; Steffen et al. 2002). The IHF protein of *E. coli* binds tightly to cognate sites represented by the consensus WATCARXXXXTTR (W is A or T; X is A, T, C or G; R is A or G; see above). There are thus two sequence elements WATCAR and TTR, often accompanied by a third A/T-rich element found upstream of WATCAR (see above). Such A/T-rich segments could have a role in modulating the assembly of the adjacent nucleoprotein complex. In contrast, a specific HU DNA binding sequence has not been detected, although HU strongly prefers various distortions in DNA such as nicks, gaps, cruciforms and phased loops (Balandina et al. 2002; Swinger and Rice 2007) and can even bind single-stranded DNA (Kamashev et al. 2008).

The crystal structure of IHF bound to DNA (Lynch et al. 2003; Rice et al. 1996; Swinger and Rice 2004) clearly shows that most of the protein contacts with DNA occur through the phosphate or sugar backbone. In fact, only three protein side-chains form hydrogen bonds with the DNA bases. It is therefore clear that IHF recognizes its cognate sites mostly through a sequence-dependent structure and flexibility of the target DNA rather than through a direct readout of the nucleotide sequence. This fact is quite counterintuitive. How can a specific sequence be

recognized indirectly through the topology of the DNA instead of its base composition? Despite many efforts to understand this, several aspects remain elusive. Although HU and IHF have a common protein fold with shared charge distributions and a notable sequence similarity (~30%), the length of the DNA binding site is longer for IHF than for HU (~35 vs ~9 bp). In addition, the bending angle induced in the DNA by IHF is generally more acute ($\geq 160^\circ$) although the length and angles for HU can be increased in vitro by varying binding conditions (Bonnefoy and Rouvière-Yaniv 1991). Despite these differences, IHF and HU are often functionally exchangeable regardless of whether the organism hosts both proteins (e.g. *E. coli*) or only one. In the microorganisms where only one protein of this type is present, such a factor is always similar to HU. In other words, unspecific facilitation (by HU-like proteins) of DNA curvature seems to be far more necessary for cell physiology than specific DNA bending. As a consequence, many bacteria have only HU, others have both HU and IHF but (as it seems) none has IHF only. IHF mutants of *E. coli* (and other IHF/HU-containing bacteria) are perfectly viable, indicating that the corresponding functions can be taken over by HU. On the contrary, mutants lacking the two HU subunits are unstable. Finally cells lacking both IHF and HU are barely viable.

What is the basis for such an unusual diverse behaviour of these two otherwise quite similar proteins (Benevides et al. 2008)? It is a well-accepted fact that residues of functional or structural significance are usually conserved within a protein family. A variation in these residues implies that some modulation or minor variation on function or structure is to be expected, and therefore the proteins bearing such changes can be thought of as a separate subfamily within the main protein family. The *Sequence Space* tool implements an algorithm (Casari et al. 1995) based on principal components analyses that facilitates the recognition of those *identity residues* that are characteristic of protein subfamilies. Such residues are called *tree-determinants* and their identification allows predictions of the protein sites responsible for the distinct function of such protein subfamily. Since -as mentioned above- protein-DNA interactions for this type of proteins are mainly indirect and therefore difficult to categorize, we have analyzed the sequence conservation of sequences in IHF and HU proteins and the presence of *tree-determinants* as an instrument for a better understanding of the role of certain amino acids in the protein structures and the consequences for their interactions with target DNA.

16.6 IHF of *Pseudomonas putida* as a New Reference for the Protein Family

Our laboratory focuses on *Pseudomonas putida* as a model organism to investigate the metabolic capabilities of soil bacteria to degrade organic compounds released by industrial activity. Because of this, we have adopted the sequences of the IHF protein of this bacterium as a reference for a model frame useful for comparisons. Although the 3D-structure of the IHF protein of *P. putida* is not available either

from X-ray diffraction or NMR studies, it shows a high sequence identity with IHF from *E. coli*, that is well characterized (Lynch et al. 2003; Rice et al. 1996; Swinger and Rice 2004). On this basis we have built a threaded structure of the heterodimer through a simple homology modeling procedure. To this end, a multiple alignment was generated after a Psi-Blast search for IHF like proteins using the sequence corresponding to both chains A and B from *P. putida*. The alignment shows a high degree of homology among all the sequences with highly conserved regions already described in the literature and thought to be important for the interaction with DNA (Fig. 16.1 and (Rice et al. 1996; Swinger and Rice 2004). A high degree of similarity/identity between the proteins of *E. coli* and *P. putida* can be clearly observed from this alignment (87% and 74% identity for chain A and B, respectively). Interestingly, such identity drops down to 45% and 55%, respectively, when one takes into account only the DNA-interacting regions of the respective proteins. For the subsequent analysis, the two chains of IHF were treated separately (one at a time), since experimental data suggest that the involvement of both chains in DNA binding is not the same. While IHF-A participates in the recognition of the WATCAR of the consensus, IHF-B seems to detect TTG (the so called H' region; (Lee et al. 1992). Models for each of the chains were generated using Swiss-PdbViewer (a protein structure homology-modeling server) with the alignments obtained previously (Guex and Peitsch 1997). The accuracy of the model structures thereby generated was analyzed for packing quality, rotamer normality, bond angles and lengths, side chain planarity, dihedral angles, etc., using the tools available at the WhatIf server (Rodriguez et al. 1998). The model proposed for IHF of *P. putida* is shown in Fig. 16.1. It is worthwhile to note that while the well known structure of the IHF-DNA complex for *E. coli* is taken as a reference (Rice et al. 1996; Swinger and Rice 2004) the numbering of the residues is corrected to reflect the predicted arrangement of the model generated for the complex IHF-DNA of *P. putida*. An analysis of the superposition of the structures for both *E. coli* and *P. putida* shows that there are no significant differences as can be inferred from the rmsd value of 0.09 for all residues obtained for both structures (model and template). In addition, the amino acid residues involved in DNA binding (and placed on those regions highly conserved) are identical. Therefore it can initially be assumed that interactions are preserved in both cases.

The structural model generated by homology methods for the integration host factor of *P. putida* exhibits the general characteristics corresponding to the nucleoid-associated bacterial proteins IHF, HU and TF1, characterized by two *beta*-arms that function as non-specific binding sites for DNA. Such *arms* are comprised of three DNA-binding HU motifs, a 3-element fingerprint for the prokaryotic integration host factor family. Within these conserved regions at the C-terminal portion of the alignment, **motif 1** spans beta-strand 1 (residues 41–56), **motif 2** encodes strands 2 and 3 (residues 59–72), and **motif 3** encompasses strand 4 and alpha-helix 3 (residues 75–89; Fig. 16.1). The charges and the distribution of polarity are maintained between IHF proteins of *E. coli* and *P. putida* (not shown). It has been speculated that protein-DNA interactions depend not just on direct recognition of base-pair, but are also significantly affected by structural properties of the DNA, such as the

features of the major and minor grooves, backbone, intrinsic curvature, hydration shells or spines, and flexibility or deformability (Steffen et al. 2002). In the case of IHF, deformation energy has been shown to be correlated with binding affinity (Aeling et al. 2006; Goodman et al. 1992). However, since not all known IHF binding sites have low deformation energy, it is possible that (i) not all IHF/DNA complexes have the same conformation, (ii) the direct recognition energetics contribute more than the indirect recognition energetics, or (iii) that other indirect effects are operative. For example, IHF binding sites often exhibit an upstream AT-rich region, which may have an associated hydration spine that contributes additionally to indirect recognition via an accommodating sequence-directed architecture. This could play a role in the regulation of the binding strength, although there are many known strong IHF-sites that lack this upstream AT-rich region, thus precluding this region from being essential (Engelhorn and Geiselmann 1998; Goodman et al. 1992).

The direct recognition interactions here taken into consideration are those within the main DNA interaction region (the consensus) and the AT-rich region. Note that the DNA consensus region for IHF binding in *Pseudomonas putida* (Valls et al. in preparation) is HAWCARnnnnWTR (being H = A, C, T not G; W = A, T; R = A, G and n = any base), that widens to an extent the one described for *E. coli* (Swinger and Rice 2004). Although this sequence is the main interacting region, our analysis reveals only five direct protein-DNA base interactions: R*59(NH) with C36(O2), R*62(NH) with G-36(N3), R*62(N) with A37(N3), R*62(NH) with T-37(O2) and the only direct interaction with the H' region R*46(NH) with T44(O2). This situation (summarized in Fig. 16.2 and Table 16.1a and b) is due to the fact that most of the interactions take place via water molecules or contacts with the DNA phosphate backbone (Vander Meulen et al. 2008). Base pairs 35 and 43 interact via water molecules with the protein, base pairs 34, 35, 36, 37, 45 through phosphate groups; and only base pairs 35, 36, 37, 44 are involved in direct interactions with protein residues. Other important interactions are those involving P64 from both chains and base pairs 37–38 and 28–29, where a proline residue is located. Protein interactions with the upstream AT-rich region must be driven mainly by structural features of the DNA as it has been previously described (Olson et al. 1998; Suzuki and Yagi 1995) and take place via the phosphate backbone as described in (Rice et al. 1996) and shown in Table 16.1a and b.

16.7 Building Family Alignments for IHF, HU, TF1

Among those proteins belonging to the IHF-like DNA-binding proteins superfamily, as described in SCOP (Structural Classification of Proteins, structural and evolutionary description of structurally characterized proteins), there is just a main family group that includes four types of proteins: IHF, HU, TF1 and H-NS. This classification is partially consistent with that of CATH (Protein Structural Classification, hierarchical domain classification of protein structures in PDB), since H-NS is not included in the same group as IHF, HU and TF1. This apparent

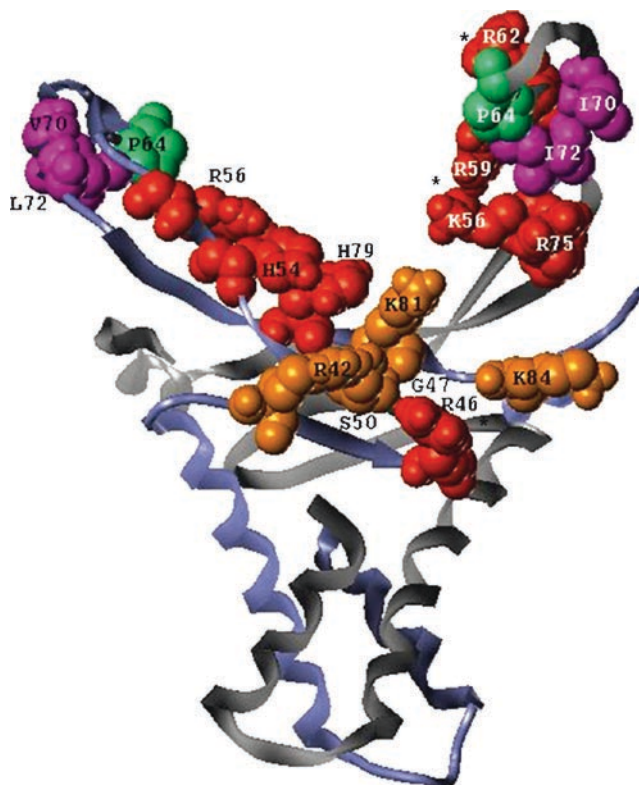


Fig. 16.2 Ribbon representation of the IHF structure (*P. putida*), showing in *space filling* mode those residues involved in interactions with DNA. The residues interacting with basepair regions 33–38 and 43–45 (according to the numbering of DNA from the *lih*f structure) are shown in red. Those interacting with bases 39–42 are orange. The bases contacting P64 are green. Finally the amino acids at positions 70 and 72, which are involved in hydrophobic interactions with DNA, are magenta. The residues indicated with an asterisk are involved in direct interactions with DNA bases

contradiction lays in the fact that one platform classifies by architecture while the other does by topology. In fact, the four proteins have identical architecture. This means the same overall shape of the domain structure as determined by the orientations of the secondary structures but ignoring the connectivity between the same secondary folding. In contrast, if one considers both the overall shape and the connectivity of the secondary structures the same proteins have two different topologies epitomized by HU and H-NS. For the sake of this Chapter we focus only on those proteins with a HU-type of topology.

Out of the various resolved structures for each of these protein types, we selected *lih*f (IHF), *lhuu* (HU), *lhue* (HU), *lb8z* (HU), *lwtu* (TF1) and *lexe* (TF1), as the more representative to generate family alignments (6, 14–18). Since, *lhue* and *lhuu* have identical sequences, only *lhuu* was selected along with *lb8z* (60% similarity between them). Similarly, since *lexe* and *lwtu* are more than 95% identical only

Table 16.1 (a) Interactions of the IHF protein with the DNA consensus sequence and the AT-rich region

Interactions via phosphate or sugar ring					
DNA		Residues involved ^a			
Region	Base	Tree-determinants		Fully conserved ^b	
		Chain A	Chain B	Chain A	Chain B
Consensus region	A34	–	HK54, R56	–	–
	T35	K56	HF79	R59	–
	C36	K56, R75	–	R59	–
	A37	R75	–	–	–
	A-41	–	R42, S50*	–	–
	C-42	–	K84	–	G47
	G45	–	R46	–	–
	A-29	–	K75	–	R59, R62
Proline insertion A-T rich region	T-19	–	K27	–	–
	T-20	–	S4	–	–
	T-21	K44	–	–	–
	A23	S46, K85*	–	–	–
	A24	K85*	–	–	–
Interactions via purine or pyrimidine bases					
Consensus region	T35	–	–	R59	–
	C36	–	–	R59	–
	G-36	–	–	R61	–
	A37	–	–	R61	–
	T-37	–	–	R61	–
	T44	–	R46	–	–
Interactions via insertion of proline					
Proline insertion	28–29	–	–	–	P64
	37–38	–	–	P64	–

(b) Summary of the interactions of the IHF protein with the DNA consensus sequence and the AT-rich region and their correlation with the degree of conservation of the residues involved as inferred from the full alignment of IHF, TF1 and HU protein families

Interactions via phosphate or sugar ring					
DNA		Residues involved ^a			
Region	Base	Tree-determinants		Fully conserved ^b	
		Chain A	Chain B	Chain A	Chain B
Consensus region	A34	–	HK54, R56 ^c	–	–
	T35	K56	HF79*	R59	–
	C36	K56, R75	–	R59	–
	A37	R75	–	–	–
	A-41	–	R42, S50*	–	–
	C-42	–	K84	–	G47
	G45	–	R46	–	–
	A-29	–	K75	–	R59, R62

(continued)

Table 16.1 (continued)

Interactions via phosphate or sugar ring					
DNA		Residues involved ^a			
Region	Base	Tree-determinants		Fully conserved ^b	
		Chain A	Chain B	Chain A	Chain B
A-T rich region	T-19	—	K27	—	—
	T-20	—	S4	—	—
	T-21	K44	—	—	—
	A23	S46, K85*	—	—	—
	A24	K85*	—	—	—
Interactions via purine or pyrimidine bases					
Consensus region	T35	—	—	R59	—
	C36	—	—	R59	—
	G-36	—	—	R62	—
	A37	—	—	R62	—
	T-37	—	—	R62	—
	T44	—	R46	—	—
Interactions via insertion of proline					
Proline	28–29	—	—	—	P64
Insertion	37–38	—	—	P64	—

^aDegree of conservation of the residues is inferred from the full alignment of IHF protein family. Residues with an asterisk (*) are considered *tree-determinants* for just $\beta\gamma$ -proteobacteria, which are not conserved among the α -proteobacteria

^bDegree of conservation of the residues is inferred from the full alignment of IHF, TF1 and HU protein families. Residues with one asterisk (*) are considered *tree-determinants* for just $\beta\gamma$ -proteobacteria, which are not conserved among the α -proteobacteria.

^cResidues in bold and underlined are *tree-determinants* for chain B of IHF that are also conserved in HU and TF1. The rest of the *tree-determinants* are just found in the corresponding chain of IHF

lexe was selected. In addition, HU and TF1 proteins were generally homodimers, while IHF is a heterodimer. Therefore, homologous searches for chain A of 1huu, 1b8z and 1exe and chains A and B of 1ihf were carried out using PsiBlast (Altschul et al. 1997). Then, the homologous sequences found for each were aligned by families using Clustal-W (Thompson et al. 1994) thereby generating an alignment for IHF-A, IHF-B, HU and TF1 subfamilies. Subsequently, the profile alignment tool of Clustal-W was used to perform consecutive alignments with IHF-A and IHF-B subfamilies, HU and TF1 subfamilies. This resulted in the generation of a final alignment for all families studied. This was used for examination of IHF-DNA interactions described for the 1ihf crystal in Rice et al. (1996) as well as for the *P. putida* model using the public server WhatIf, a suite of programs designed to modify and check the validity of pdb entries (Rodriguez et al. 1998). All these data were the basis for searching conserved and *tree-determinant* residues by using the SequenceSpace package (Casari et al. 1995) on the alignment corresponding to both chains of IHF and on the full alignment including IHF/HU/TF1. The implemented algorithm clusters the aligned protein sequences and projects the sequence residues

into the protein space, so that groups of residues specific for the clusters are revealed, allowing the identification of amino acids that are predicted to be directly involved in protein function.

16.7.1 Comparative Analysis of IHF Proteins

Analysis of the protein IHF family alignments results in 45–50% identity for IHF-A and 45–50% for IHF-B. If only the DNA-binding regions are considered, the identity increases to 50–55% for IHF-A and ~55% for IHF-B (Table 16.3). The alignment comprising both chains resulted in 30–35% for IHF-A/IHF-B, 25–30% and ~40% for the full alignment or just the DNA-binding region, respectively. The evolutionary tree clearly reflects the sequence differences between both chains of IHF and also the differences within each chain between α -proteobacteria and β - γ -proteobacteria (Gram-negative), γ -proteobacteria being paraphyletic (the group contains its most recent common ancestor but does not contain all the descendants of that ancestor) with respect to β -proteobacteria (as shown in Fig. 16.3). The resulting alignment for all the sequences with less than 80% of homology is shown in Fig. 16.4. This alignment reveals certain features for instance, a number of residues fully conserved for all sequences that is placed mainly at the DNA-binding region

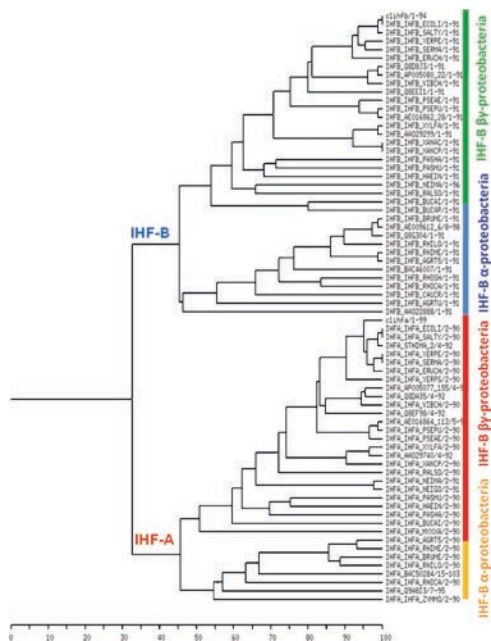


Fig. 16.3 Alignment of IHF-A and IHF-B homologous sequences and distribution among different groups of proteobacteria

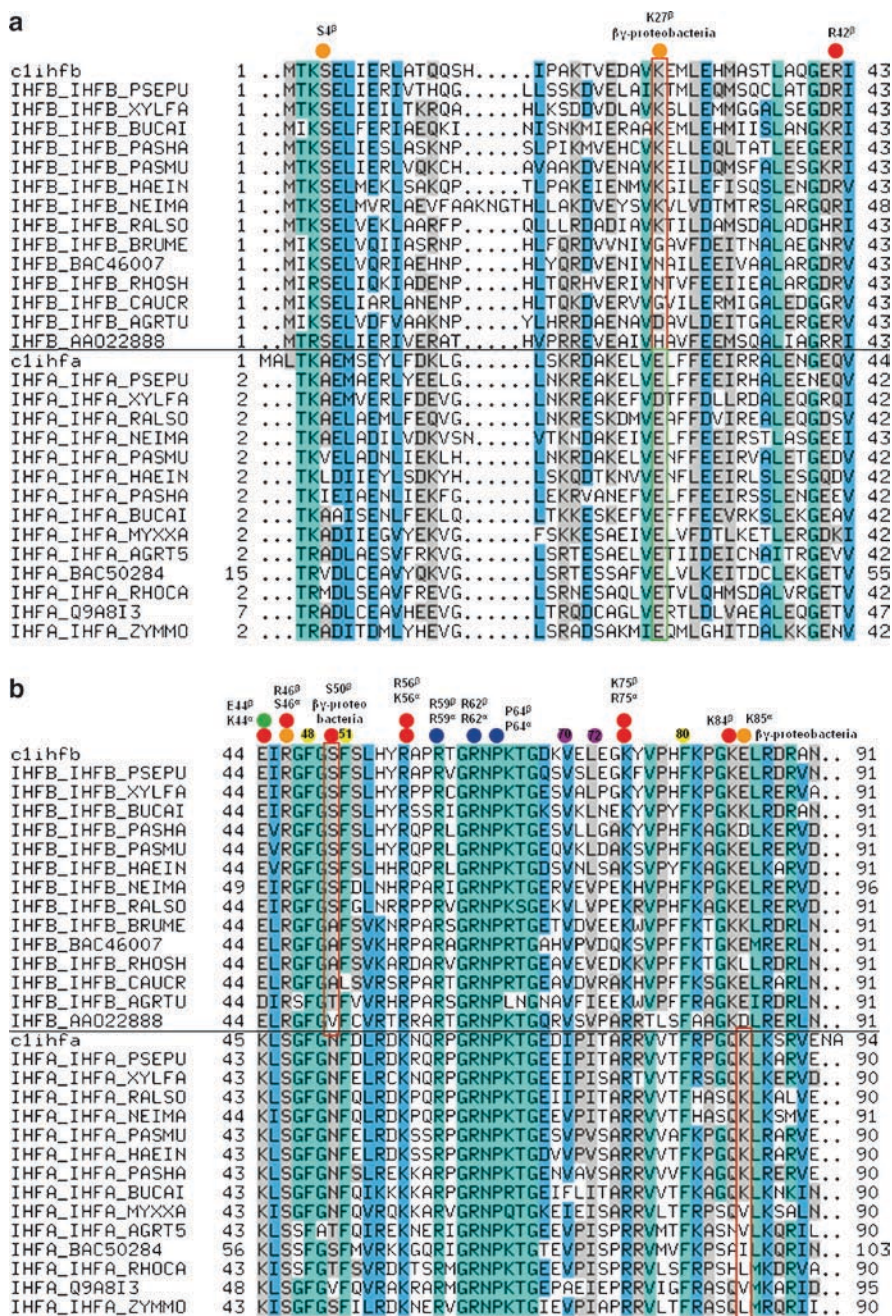


Fig. 16.4 Profile alignment for IHF-A and IHF-B homologous sequences (80% of non-redundancy). Note the inclusion of the IHF-A of *P. putida* (83% homologous to IHF-A of *E. coli*). Several key residues for each chain are marked

(such as F48, F51 and F80). There are also residues in the segment 59–64 that are fully conserved in one of the chains but not in the other. These residues are named *tree-determinants* (see above) since they define the point where the protein tree branches between the two chains, which is the case for residue 46: this site is an arginine for chain-B, but it is a serine for chain-A. Similarly, residue 44 is glutamic acid in chain B and lysine in chain A.

Application of *SequenceSpace* tools provides a more complete analysis of the degree of conservation of the protein residues (Casari et al. 1995). This method uses a vectorial representation of each protein sequence as a point in a multidimensional space along with multi-variant statistics and principal component analysis, to reduce the number of dimensions. This depiction allows us to define clusters of proteins according to specific properties by choosing the appropriate axes defined by the highest corresponding *eigenvalues* (also known as proper values). Moreover, this type of representation also permits to project the individual residues on the same axes, and thus trace the positions conserved in the subfamilies under examination. The main advantage of this method is the possibility of predicting which residues may be responsible for the specific characteristics of each protein subfamily or group of subfamilies. The two-dimensional projection of sequence vectors on the plane defined by the axes corresponding to eigenvalues 2 and 3 revealed the clustering of protein subfamilies (chain A, chain B differentiated into α -proteobacteria and $\beta\gamma$ -proteobacteria) according to the DNA-interaction properties (Fig. 16.5a). The projection of the individual amino acid residues on the same plane (Fig. 16.5b) reveals that the amino acids responsible for this segregation might be responsible for the differences between chains A and B in their way of interaction with DNA (Tables 16.2 and 16.3). Note that residues fully conserved along all sequences are those placed at the tip of the two arms, and the 3 phenylalanines at the core of the

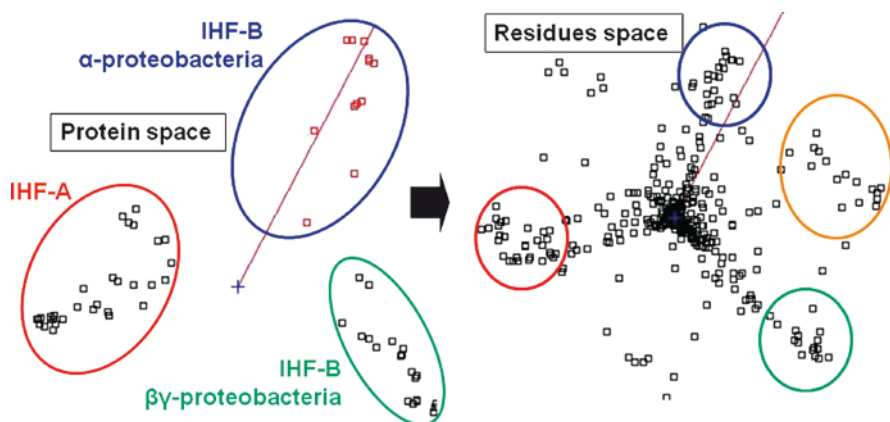


Fig. 16.5 Sequence space and residues space projected onto the plane defined by eigenvalues 2 and 3 for the alignment by profiles of IHF-A and IHF-B homologous sequences. Characteristic residues for IHF-A (red), IHF-B (orange), chain B of α -proteobacteria (blue) and chain B of $\beta\gamma$ -proteobacteria in (green) are indicated

Table 16.2 Degree of conservation of the protein residues as inferred from the full alignment of IHF protein family chains A and B

(A) Conservation degree	Residues ^a	
Fully conserved	55 F (48), 58 F (51), <u>66R</u> (59), 68 G (61), <u>69R</u> (62), 70 N (63), <u>71P</u> (64), 74 G (67), 87 F (80)	
Highly conserved	5 KR (3), 7 DE (5), 12 VIL (10), 33 VI (26), 44 LI (37), 50 VI (43), 52 VIL (45), 54 G (47), 72 KR (65), 73 T (66), <u>77VI</u> (70), <u>79VIL</u> (72), 84 V (77), 93 L (86), 97 VIL (90)	
Tree determinants	IHF-A	all34 ED (27), 51 K (44), 53 S (46), <u>61R</u> (54), <u>63K</u> (56), 79 I (72), <u>82R</u> (75), <u>94K</u> (87)
		$\beta\gamma$ -proteobact.30 K (23), 36 F (29), 41 R (34), 57 N (50), 67 P (60), 85 V (78), 91 Q (84)), 92 K (85)
	IHF-B	all6 S (4), 29 VI (22), 30 ED (23), <u>49R</u> (42), 51 E (44), <u>53R</u> (46), <u>61HK</u> (54), 63 R (56), 79 LV (72), <u>82K</u> (75), 85 P (78), 91 K (84), 94 R (87)
		α -proteobact.4 I (2), 16 N (14), 17 P (15), 57 A (50), 86 F (79), 97 L (90)
		$\beta\gamma$ -proteobact.34 K (27), <u>57S</u> (50), <u>86HY</u> (79)
(B) Conservation degree	Residues ^b	
Fully conserved	55 F (48), 58 F (51), 66 R (59), 68 G (61), 69 R (62), 70 N (63), 71 P (64), 74 G (67), 87 F (80)	
Highly conserved	5 KR (3), 7 DE (5), 12 VIL (10), 33 VI (26), 44 LI (37), 50 VI (43), 52 VIL (45), <u>54G</u> (47), 72 KR (65), 73 T (66), 77 VI (70), 79 VIL (72), 84 V (77), 93 L (86), 97 VIL (90)	
Tree determinants	IHF-A	all34 ED (27), 51 K (44), <u>53S</u> (46), 61 R (54), 63K (56), 79 I (72), 82 R (75), 94 K (87)
		$\beta\gamma$ -proteobact.30 K (23), 36 F (29), 41 R (34), <u>57N</u> (50), 67 P (60), 85 V (78), 91 Q (84)), <u>92K</u> (85)
	IHF-B	all <u>6S</u> (4), 29 VI (22), 30 ED (23), 49 R (42), 51 E (44), 53 R (46), 61 HK (54), 63R (56), 79 LV (72), 82 K (75), 85 P (78), 91 K (84), 94 R (87)
		α -proteobact.4 I (2), 16 N (14), 17 P (15), 57 A (50), 86 F (79), 97 L (90)
		$\beta\gamma$ -proteobact.34 K (27), <u>57S</u> (50), 86 HY (79)

^aResidue positions correspond to alignment sites. The position numbering corresponding to the IHF of *P. putida* are indicated in parenthesis. Residues involved in interactions with DNA are underlined. Residues involved in interactions with the AT-rich region of DNA are in italics.

^bResidues involved in interactions with DNA are in bold. Residues interacting with the AT-rich region of DNA are underlined. Residues in italics are those tree-determinants lost after inclusion of HU and TF1 families in the alignment.

Table 16.3 Degree of sequence homology between the pairs of protein families resulting from profile alignments

Whole sequence ^a		IHF		HU	TF1
		IHF-A	IHF-B		
IHF	IHF-A	45–50%	25–30%	30–35%	25–30%
	IHF-B	25–30%	45–50%		
HU		30–35%		45–50%	35–40%
TF1		25–30%		35–40%	35–30%
DNA-binding region ^a					
	IHF		HU	TF1	IHF-A
IHF-B	IHF	IHF-A			50–55%
~40%		40–45%	35–40%	IHF-B	~40%
~55%	HU			40–45%	
45–50%	35–40%	TF1		35–40%	
35–40%	35–30%				

^aSee *comparative analysis of IHF proteins* in text for explanation

protein-dimer (Fig. 16.6). Tables 16.2a and b also enumerate a group of highly conserved residues, i.e. positions with conservative mutations. Most residues in this group are hydrophobic and are forming the core of the protein-dimer, with exception of positions 3, 5 and 65. Therefore, the majority of the *fully-conserved* and *highly-conserved* residues are those defining the tip of the two arms (those which initiate the wrapping of DNA around the protein) and the stable hydrophobic core of the protein. Table 16.2a and b also shows those residues characteristics for each of the two chains of IHF proteins. Therefore it is possible to find along all these residues some of the key amino acids for interactions with DNA.

Analysis of the degree of conservation of the IHF sites involved in interactions with DNA reveals that most of such amino acids belong to the *fully-conserved* or to the *tree-determinant* groups residues as shown in Fig. 16.7. To illustrate how DNA recognition takes place, Fig. 16.8 shows the distribution of these residues by groups, within the spatial 3D-structure of the protein. The figure shows how the amino acids interacting with DNA at the tips are *fully-conserved*, whereas those that lay along DNA as well as those interacting with the AT-rich DNA region are *tree-determinants*. A more detailed analysis reveals that two out of the three residues involved in direct interactions (R59, R62 plus P64) with DNA bases are *fully-conserved* residues, the third one (R46) being characteristic of the IHF chain B. Moreover, while chain B is interacting with the two regions of the DNA-consensus, chain A is not interacting with the H'-region. There are more differential features between chains A and B. In chain A half of the residues involved in interactions with DNA are *tree-determinants*, while the other half are *fully-conserved* amino acids. In contrast, most DNA-touching counterparts for chain B are *tree-determinants* (Table 16.1a and b). P^α64 gets inserted between bp 37–38 (i.e. in the DNA-consensus region), while P^β64 is inserted between bp 28–29. E^β44, that has been described as a key residue to keep in position residues R^β46 and R^β42 to properly

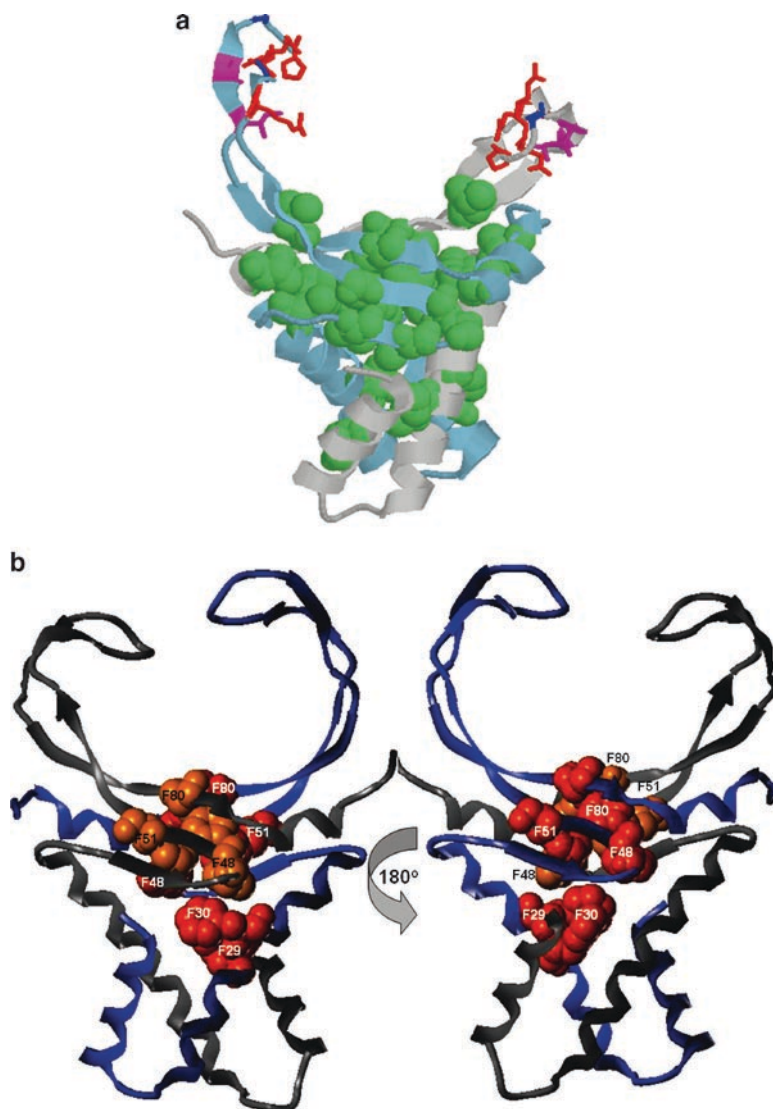


Fig. 16.6 Structural model of IHF of *Pseudomonas putida* interacting with its target DNA. (a) Conserved non-polar residues (as inferred from sequence space analysis above) are labeled in green. Residues IV70 and IVL72 involved in hydrophobic interactions with DNA are painted in magenta; G61 and G67 are in blue, while P64, R59, R62 are in red. (b) Highlighting of key phenylalanine residues

interact with DNA bp 44 by making salt bridges with the guanidinium moieties is conserved along all IHF-B. E^β44 it thus a *tree-determinant* residue characteristic of the IHF chain B. Interaction with the AT-rich region of DNA takes place via *tree-determinant* residues of both chains and it is more specific of each of the proteobacterial subgroups.

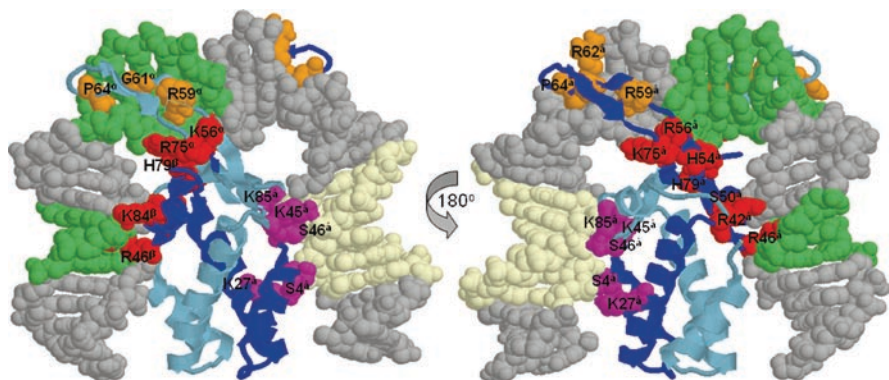


Fig. 16.7 Detail of the 3D model for IHF-DNA interaction in *P. putida*. The figure shows those residues involved in interactions with the different regions of the target DNA. The DNA consensus region is shown in green, while the AT-rich region of DNA is painted light-yellow. Those protein residues which are *fully-conserved* and involved in interactions with DNA are represented in with orange. Finally, *tree-determinant* residues involved in interactions with DNA are in red, while *tree-determinants* concerned with interactions with the AT-rich region of DNA are magenta

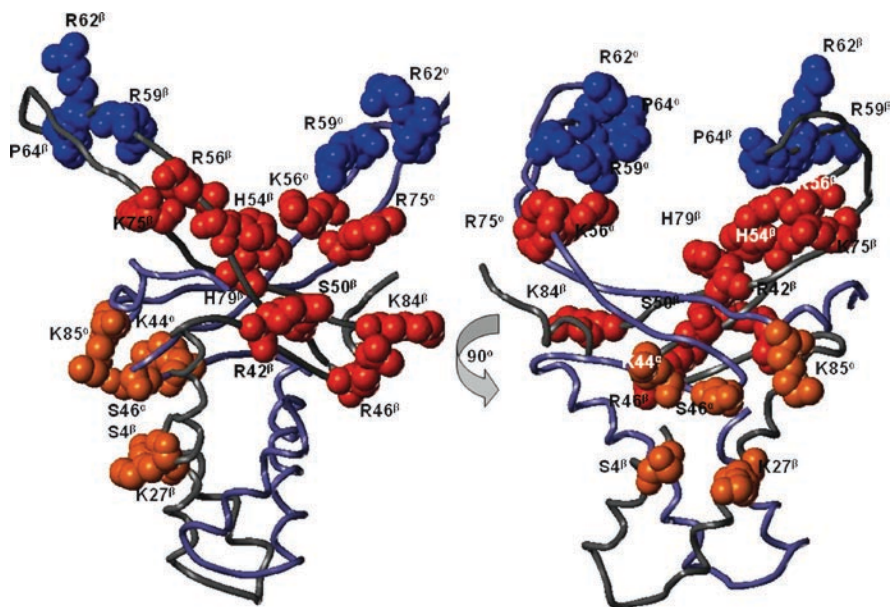


Fig. 16.8 Architecture of the distribution of IHF residues involved in interactions with DNA and coloured by the degree of conservation within the family. *Tree-determinant* residues involved in interactions with DNA are shown in red-orange; those fully-conserved also involved in interactions with DNA are in blue; *tree-determinants* involved in interactions with the AT-rich region of DNA are in orange

The emerging picture of all these analyses is that IHF-DNA interactions are initially driven by the interaction of the tips of the arms and then secured by specific interactions with the *tree-determinant* residues.

16.7.2 What We Learn from Comparing IHF with HU and TF1

HU is another key member of the prokaryotic DNA-binding, chromatin-associated protein repertoire. The protein shares with IHF many structurally important features such as the distribution of positive charge and the presence of prolines at the tips of the arms. It would be reasonable to expect that the HU-DNA complex would be very similar to that of IHF. However, as mentioned above and explained in detail in Chapter 16, HU binds DNA in a non-specific manner (Bonnefoy and Rouvière-Yaniv 1991; Bonnefoy and Rouvière-Yaniv 1992; Kamashev et al. 1999). It has been suggested that the sequence specificity of IHF probably results from the sum of a number of small differences between the two proteins. Some of the described differences between HU and IHF proteins include the replacement of R^β46 in HU or TF1 proteins, the axial displacement of helix 1 in IHF-A vs. IHF-B or HU. In this case, the asymmetry of interactions between DNA and IHF may drive the two arms to make different interactions with the minor groove (Lynch et al. 2003; Rice 1997; Rice et al. 1996; Swinger et al. 2003).

Alignment of each of the homologous sequences found for HU and TF1 resulted in identity values of 35–40% for TF1 and 45–50% for HU. Unlike the case of the two chains of IHF discussed above, these values did not change when only the DNA-binding region was considered. When these two families (HU and TF1) were compared with both chains of IHF, some interesting features were revealed. First, despite the difference in DNA-binding specificity between HU and TF1 (non-specific binding vs. specific binding, respectively) their sequences exhibit high homology among themselves and with IHF (35–40% identity). The similarity is largest in the three regions known as the DNA-binding HU motif. Another result of the comparison is that HU is more closely related to IHF in terms of sequence than TF1. However IHF and TF1 share their ability to bind specific DNA sequences. Although TF1 is a homodimer (like many HU proteins), TF1 does recognize specific DNA sites within the phage genome, and produces a hydroxyl radical footprint quite similar to that of IHF. Therefore, there seems to be a structural and functional continuum between the three proteins. This can be inferred from the alignments among families that resulted in 35–40% for HU/TF1, 30–35% for IHF-A/IHF-B, 30–35% HU/IHF (more related to IHF-B than IHF-A), 25–30% TF1/IHF and 25–30% for the full alignment. These percentages changed again for those in which IHF is involved, resulting in ~40% for IHF, 40–45% for HU/IHF and 35–40% for TF1/IHF (more related to chain B too) and 30–35% for HU/TF1/IHF.

When compared all together IHF, HU and TF1, one can observe that HU and TF1 form a bona fide protein family, while IHF-A and IHF-B are two branches that clearly are separated from HU/TF1. However, when one just takes into account the

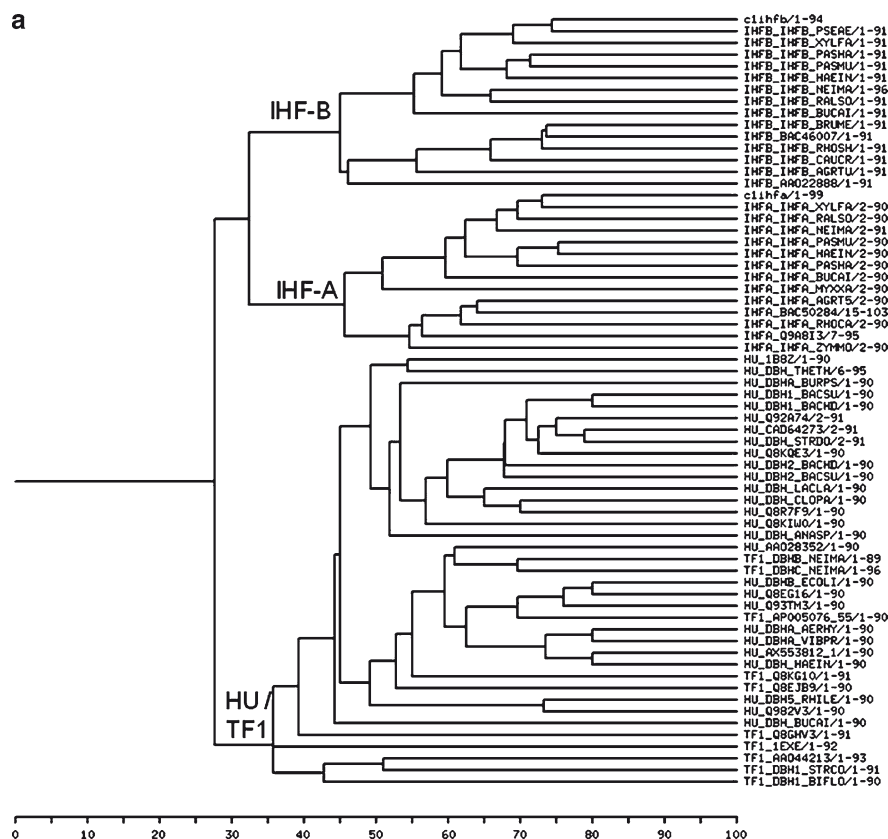


Fig. 16.9 Tree diagram representing protein family distribution for IHF chains A and B vs HU/TF1 family. The sequences examined are those with degree of homology lower than 80%. (**a**) Analysis of the whole protein.

DNA-binding regions, one can see how IHF-A is clearly a branch that diverges from the one including IHF-B and HU/TF1 (Fig. 16.9). Therefore, it is possible that IHF-B and HU/TF1 are more closely related evolutionarily than IHF-A, while the non-interacting regions of IHF-A and IHF-B may have evolved simultaneously.

Analysis of the family alignment for IHF/HU/TF1 sequences reveals a high degree of homology along those protein fragments involved in interaction with DNA (the DNA binding HU motifs mentioned above). These residues considered as *fully-conserved* for the IHF family remain as highly conserved when HU and TF1 families are considered (shown in blue in Fig. 16.10). Of the eight phenylalanine residues forming part of the hydrophobic core ($F^{\alpha 29}$, $F^{\alpha 30}$, $F^{\alpha 48}$, $F^{\beta 48}$, $F^{\alpha 51}$, $F^{\beta 51}$, $F^{\alpha 80}$, $F^{\beta 80}$) only $F^{\alpha 29}$, $F^{\alpha 30}$ are specific of IHF-A β -proteobacteria while the other are highly conserved in all IHF/HU/TF1. Therefore, since these residues are located at the core of the dimer and at the tip of the DNA-binding arms, one can assume that the overall structure of the core dimer and the interactions with DNA

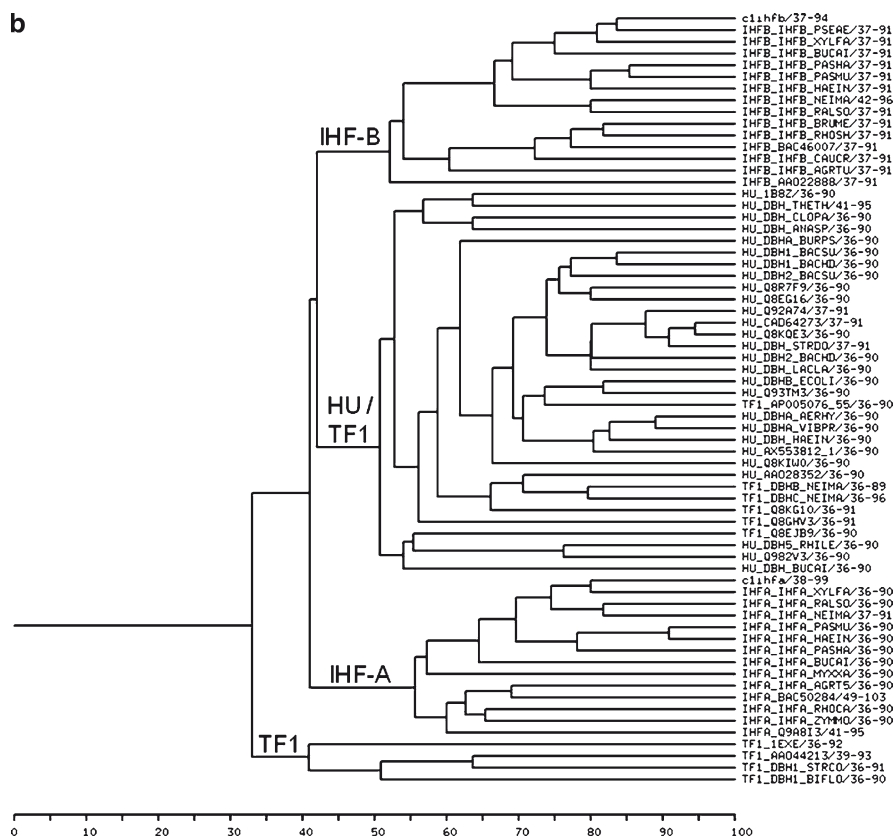


Fig. 16.9 (continued) (b) Analysis of DNA-binding regions only

through the tip of the arms remains the same for HU and TF1 groups. Unlike IHF sequences, endowed with a relatively large number of residues considered as *tree-determinants*, only the N2, K19, Q65 can be considered as such in HU and TF1.

In a subsequent step, we used the *SequenceSpace* algorithm (Casari et al. 1995) to perform an in-depth analysis of those key residues of IHF that match or not the corresponding sequences of the HU and TF1 protein groups. As indicated before, Table 16.2 shows how those *fully-conserved* or *highly conserved* residues (positions with conservative mutations) remain as such. R59 and R62 are found among the fully-conserved residues located at the tip of the arms and involved in direct interactions with DNA bases. Other conserved residues at positions 70 and 72 (also at the tip of the arms) are involved in hydrophobic interactions with DNA. Finally, P64 is responsible of generating the kink in the DNA. These residues are conserved for all HU and TF1 sequences, suggesting a putative similar interaction with DNA at this level. On the contrary, R^B46 (the third residue involved in direct interactions with DNA bases) and the E^B44 site (bringing about H-bond formation to stabilize interaction between R^B46 and T44 of DNA) are specific for chain B, and are lost in

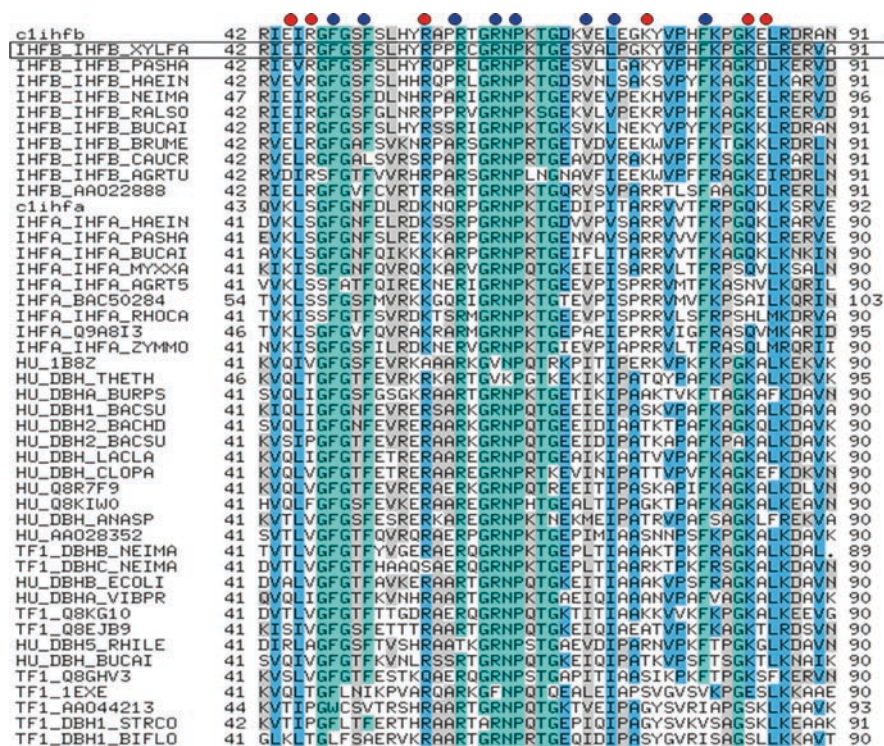


Fig. 16.10 Profile alignment for IHF-A, IHF-B and HU/TF1 homologous segments. The sequences examined are those with $\geq 80\%$ of non-redundancy

HU or TF1 sequences (Table 16.1b). These interactions include the H'-region of the DNA consensus. This is in contrast with those of R59 and R62, which involve the first region of the DNA-consensus (33–38 bps). In addition the remaining residues of IHF involved in interactions with DNA (all of which are chain-specific) are not shared by either HU or TF1 sequences i.e., they are specific of IHF. As an exception (all of them at chain B), residues R^B56 and K^B84 (interacting with A34 and C42 of the first region of the consensus) are present also in all the HU and TF1 sequences. Finally, residue S^B4, (involved in interactions with the AT-rich region of DNA), is present in some (but not all) HU and TF1 sequences (Table 16.1b). In summary, the number of *tree-determinants*, which are lost after including all sequences from HU and TF1 is higher for IHF-B than for IHF-A. This supports the idea of a closer evolutionary relationship between IHF-B and HU and TF1 groups.

16.8 Conclusion

The comparative analysis of the diverse regions of IHF, HU and TF1 that interact with DNA reveals interesting insights into the way in which IHF specifically binds and bends target sequences. **First**, protein interactions with the AT-rich region in the DNA are driven by *tree-determinant* residues, which are specific for each chain of IHF and are not generally found in the sequences of HU or TF1. Yet, these interactions do vary as a function of the type of proteobacteria involved. **Second**, those fully-conserved amino acids involved in direct interactions with the DNA bases are placed at the tip of the protein dimer arms and they are maintained in HU and TF1 proteins. Most *tree-determinant* residues involved in interactions with DNA (including the fourth R^B46) are specific of the IHF sequences, with the exceptions indicated above. Therefore the interactions that fix the specificity are those that take place at the first region of the consensus (Fig. 16.11). The picture that emerges from this study is that indirect recognition of target sequences is followed by direct recognition and not the contrary. In other words, the high specificity of IHF binding to DNA is achieved despite a relaxed sequence dependency. The analysis presented above supports the concept that the *tree-determinants* only found in IHF proteins

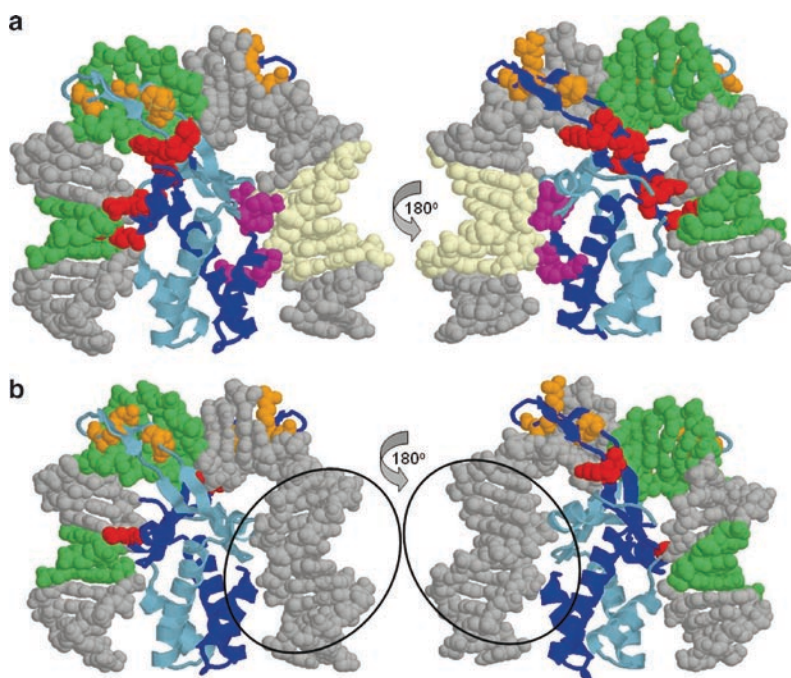


Fig. 16.11 Comparison of protein-DNA binding modes between IHF (*top*) and HU/TF1 (*bottom*) families, using as a reference the IHF-DNA interaction scheme. *Fully-conserved* residues are shown in orange, *tree-determinant* residues in red, and *tree-determinant* residues involved in interactions with AT-rich region in magenta (note their absence in the case of HU/TF1).

(but not in HU) are the ones that endow specificity to the recognition of a distinct target DNA sequence. The absence of such *tree-determinants* makes HU proteins unable to fulfill the same action. Furthermore, the maintenance of the fully conserved residues that bring about the bending of the DNA explains why HU/TFI and IHF can be partially replaced by each other. But the binding accuracy is endowed only by the *tree-determinants* that are specific for each chain of IHF. The H' region of the consensus and the AT-rich tract thus discriminates between IHF and HU, as the sequence conservation for residues involved in interactions with the AT-sequence are not conserved in HU or TFI. It will be interesting to follow in the next few years whether these predictions, largely based on sequence comparisons, are supported by biochemical experimental data from assays to probe DNA-protein interactions (Vivas et al. 2008), including single-molecule approaches.

Acknowledgments The work made in Authors' Laboratory was supported by research grants of the Ministry of Science and Innovation, by contracts of the 6th Framework Programme of the EU and by Funds of the Autonomous Community of Madrid.

References

- Abril MA, Ramos JL (1993) Physical organization of the upper pathway operon promoter of the *Pseudomonas* TOL plasmid. Sequence and positional requirements for XylR-dependent activation of transcription. *Mol Gen Genet* 239:281–288
- Aeling KA, Opel ML, Steffen NR, Tretyachenko-Ladokhina V, Hatfield GW, Lathrop RH, Senear DF (2006) Indirect recognition in sequence-specific DNA binding by *Escherichia coli* integration host factor: the role of DNA deformation energy. *J Biol Chem* 281:39236–39248
- Aeling KA, Steffen NR, Johnson M, Hatfield GW, Lathrop RH, Senear DF (2007) DNA deformation energy as an indirect recognition mechanism in protein-DNA interactions. *IEEE/ACM Trans Comput Biol Bioinform* 4:117–125
- Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181:6361–6370
- Ali BM, Amit R, Braslavsky I, Oppenheim AB, Gileadi O, Stavans J (2001) Compaction of single DNA molecules induced by binding of integration host factor (IHF). *Proc Natl Acad Sci USA* 98:10658–10663
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Balandina A, Kamashev D, Rouvière-Yaniv J (2002) The bacterial histone-like protein HU specifically recognizes similar structures in all nucleic acids. DNA, RNA, and their hybrids. *J Biol Chem* 277:27622–27628
- Benevides JM, Danahy J, Kawakami J, Thomas GJ Jr (2008) Mechanisms of specific and nonspecific binding of architectural proteins in prokaryotic gene regulation. *Biochemistry* 47:3855–3862
- Bertoni G, Fujita N, Ishihama A, de Lorenzo V (1998) Active recruitment of sigma54-RNA polymerase to the Pu promoter of *Pseudomonas putida*: role of IHF and alphaCTD. *EMBO J* 17:5120–5128
- Bonnefoy E, Rouvière-Yaniv J (1991) HU and IHF, two homologous histone-like proteins of *Escherichia coli*, form different protein-DNA complexes with short DNA fragments. *EMBO J* 10:687–696

- Bonnefoy E, Rouvière-Yaniv J (1992) HU, the major histone-like protein of *E. coli*, modulates the binding of IHF to *oriC*. EMBO J 11:4489–4496
- Calb R, Davidovitch A, Koby S, Giladi H, Goldenberg D, Margalit H, Holtel A, Timmis K, Sanchez-Romero JM, de Lorenzo V, Oppenheim AB (1996) Structure and function of the *Pseudomonas putida* integration host factor. J Bacteriol 178:6319–6326
- Casari G, Sander C, Valencia A (1995) A method to predict functional residues in proteins. Nat Struct Biol 2:171–178
- Claret L, Rouvière-Yaniv J (1997) Variation in HU composition during growth of *Escherichia coli*: the heterodimer is required for long term survival. J Mol Biol 273:93–104
- de Lorenzo V, Herrero M, Metzke M, Timmis KN (1991) An upstream XylR- and IHF-induced nucleoprotein complex regulates the sigma 54-dependent Pu promoter of TOL plasmid. EMBO J 10:1159–1167
- Delic-Attree I, Toussaint B, Vignais PM (1995) Cloning and sequence analyses of the genes coding for the integration host factor (IHF) and HU proteins of *Pseudomonas aeruginosa*. Gene 154:61–64
- Delic-Attree I, Toussaint B, Froger A, Willison JC, Vignais PM (1996) Isolation of an IHF-deficient mutant of a *Pseudomonas aeruginosa* mucoid isolate and evaluation of the role of IHF in *algD* gene expression. Microbiology 142:2785–2793
- Ditto MD, Roberts D, Weisberg RA (1994) Growth phase variation of integration host factor level in *Escherichia coli*. J Bacteriol 176:3738–3748
- Dixit S, Singh-Zocchi M, Hanne J, Zocchi G (2005) Mechanics of binding of a single integration-host-factor protein to DNA. Phys Rev Lett 94:118101
- dos Santos MT, Rodrigues PS (2005) A genomic-scale search for regulatory binding sites in the integration host factor regulon of *Escherichia coli* K12. Genet Mol Res 4:783–789
- Engelhorn M, Geiselmann J (1998) Maximal transcriptional activation by the IHF protein of *Escherichia coli* depends on optimal DNA bending by the activator. Mol Microbiol 30:431–441
- Freundlich M, Ramani N, Mathew E, Sirko A, Tsui P (1992) The role of integration host factor in gene expression in *Escherichia coli*. Mol Microbiol 6:2557–2563
- Friedman DI, Weglenska A, Jacob B, Sirko A, Pratt TS, Steiner T, Feldman LS, Walker KA, Osuna R (1988) Integration host factor: a protein for all reasons. Cell 55:545–554
- Giladi H, Koby S, Prag G, Engelhorn M, Geiselmann J, Oppenheim AB (1998) Participation of IHF and a distant UP element in the stimulation of the phage lambda PL promoter. Mol Microbiol 30:443–451
- Gober JW, Shapiro L (1990) Integration host factor is required for the activation of developmentally regulated genes in *Caulobacter*. Genes Dev 4:1494–1504
- Goodman SD, Nicholson SC, Nash HA (1992) Deformation of DNA during site-specific recombination of bacteriophage lambda: replacement of IHF protein by HU protein or sequence-directed bends. Proc Natl Acad Sci USA 89:11910–11914
- Goodrich JA, Schwartz ML, McClure WR (1990) Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor (IHF). Nucleic Acids Res 18:4993–5000
- Goosen S, van de Putte P (1995) The regulation of transcription initiation by integration host factor. Mol Microbiol 16:1–7
- Grainger DC, Hurd D, Goldberg MD, Busby SJW (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. Nucleic Acids Res 34:4642–4652
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723
- Hales LM, Gumpert RI, Gardner JF (1996) Examining the contribution of a dA + dT element to the conformation of *Escherichia coli* integration host factor-DNA complexes. Nucleic Acids Res 24:1780–1786

- Haluzi H, Goitein D, Koby S, Mendelson I, Teff D, Mengeritsky G, Giladi H, Oppenheim AB (1991) Genes coding for integration host factor are conserved in Gram-negative bacteria. *J Bacteriol* 173:6297–6299
- Hoover TR, Santero E, Porter S, Kustu S (1990) The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. *Cell* 63:11–22
- Kamashev D, Balandina A, Rouvière-Yaniv J (1999) The binding motif recognized by HU on both nicked and cruciform DNA. *EMBO J* 18:5434–5444
- Kamashev D, Balandina A, Mazur AK, Arimondo PB, Rouvière-Yaniv J (2008) HU binds and folds single-stranded DNA. *Nucleic Acids Res* 36:1026–1036
- Koh J, Saecker RM, Record MT (2008) DNA binding mode transitions of *Escherichia coli* HU (alpha-beta): evidence for formation of a bent DNA-protein complex on intact, linear duplex DNA. *J Mol Biol* 383:324–346
- Kuznetsov SV, Sugimura S, Vivas P, Crothers DM, Ansari A (2006) Direct observation of DNA bending/unbending kinetics in complex with DNA-bending protein IHF. *Proc Natl Acad Sci USA* 103:18515–18520
- Lee EC, MacWilliams MP, Gumpert RI, Gardner JF (1991) Genetic analysis of *Escherichia coli* integration host factor interactions with its bacteriophage lambda H' recognition site. *J Bacteriol* 173:609–617
- Lee EC, Hales LM, Gumpert RI, Gardner JF (1992) The isolation and characterization of mutants of the integration host factor (IHF) of *Escherichia coli* with altered, expanded DNA-binding specificities. *EMBO J* 11:305–313
- Lynch TW, Read EK, Mattis AN, Gardner JF, Rice PA (2003) Integration host factor: putting a twist on protein-DNA recognition. *J Mol Biol* 330:493–502
- Nash HA (1996) The HU and IHF proteins: accessory factors for complex protein-DNA assemblies. In: Lin EEC, Lynch AS (eds) *Regulation of gene expression in Escherichia coli*. R.G. Landes Company, Austin, TX, pp 149–179
- Nash HA, Robertson CA (1981) Purification and properties of the *Escherichia coli* protein factor required for lambda integrative recombination. *J Biol Chem* 256:9246–9253
- Ochman H, Davalos LM (2006) The nature and dynamics of bacterial genomes. *Science* 311:1730–1733
- Olson WK, Gorin AA, Lu XJ, Hock LM, Zhurkin VB (1998) DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci USA* 95:11163–11168
- Painbeni E, Caroff M, Rouvière-Yaniv J (1997) Alterations of the outer membrane composition in *Escherichia coli* lacking the histone-like protein HU. *Proc Natl Acad Sci USA* 94:6712–6717
- Parekh BS, Hatfield GW (1996) Transcriptional activation by protein-induced DNA bending: evidence for a DNA structural transmission model. *Proc Natl Acad Sci USA* 93:1173–1177
- Pearl FM, Martin N, Bray JE, Buchan DW, Harrison AP, Lee D, Reeves GA, Shepherd AJ, Sillitoe I, Todd AE, Thornton JM, Orengo CA (2001) A rapid classification protocol for the CATH Domain Database to support structural genomics. *Nucleic Acids Res* 29:223–227
- Pérez-Martín J, de Lorenzo V (1996a) In vitro activities of an N-terminal truncated form of XylR, a σ 54-dependent transcriptional activator of *Pseudomonas putida*. *J Mol Biol* 258:575–587
- Pérez-Martín J, de Lorenzo V (1996b) Physical and functional analysis of the prokaryotic enhancer of the σ 54-promoters of the TOL plasmid of *Pseudomonas putida*. *J Mol Biol* 258:562–574
- Pérez-Martín J, Timmis KN, de Lorenzo V (1994) Co-regulation by bent DNA. Functional substitutions of the integration host factor site at σ 54-dependent promoter Pu of the upper-TOL operon by intrinsically curved sequences. *J Biol Chem* 269:22657–22662
- Rice PA (1997) Making DNA do a U-turn: IHF and related proteins. *Curr Opin Struct Biol* 7:86–93
- Rice PA, Yang S, Mizuuchi K, Nash HA (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* 87:1295–1306

- Rodriguez R, Chinae G, Lopez N, Pons T, Vriend G (1998) Homology modeling, model and software evaluation: three related resources. *Bioinformatics* 14:523–528
- Schmid MB (1990) More than just “histone-like” proteins. *Cell* 63:451–453
- Senear DF, Tretyachenko-Ladokhina V, Opel ML, Aeling KA, Hatfield GW, Franklin LM, Darlington RC, Alexander Ross JB (2007) Pressure dissociation of integration host factor-DNA complexes reveals flexibility-dependent structural variation at the protein-DNA interface. *Nucleic Acids Res* 35:1761–1772
- Steffen NR, Murphy SD, Toller L, Hatfield GW, Lathrop RH (2002) DNA sequence and structure: direct and indirect recognition in protein-DNA binding. *Bioinformatics* 18:S22–S30
- Stonehouse E, Kovacicova G, Taylor RK, Skorupski K (2008) Integration host factor positively regulates virulence gene expression in *Vibrio cholerae*. *J Bacteriol* 190:4736–4748
- Suzuki M, Yagi N (1995) Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res* 23:2083–2091
- Swinger KK, Rice PA (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol* 14:28–35
- Swinger KK, Rice PA (2007) Structure-based analysis of HU-DNA binding. *J Mol Biol* 365:1005–1016
- Swinger KK, Lemberg KM, Zhang Y, Rice PA (2003) Flexible DNA bending in HU-DNA cocystal structures. *EMBO J* 22:3749–3760
- Tanaka I, Appelt K, Dijk J, White SW, Wilson KS (1984) 3-A resolution structure of a protein with histone-like properties in prokaryotes. *Nature* 310:376–381
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Ussery D, Larsen TS, Wilkes KT, Friis C, Worning P, Krogh A, Brunak S (2001) Genome organisation and chromatin structure in *Escherichia coli*. *Biochimie* 83:201–212
- Valls M, Buckle M, de Lorenzo V (2002) In vivo UV laser footprinting of the *Pseudomonas putida* σ 54Pu promoter reveals that integration host factor couples transcriptional activity to growth phase. *J Biol Chem* 277:2169–2175
- Vander Meulen KA, Saecker RM, Record MT Jr (2008) Formation of a wrapped DNA-protein interface: experimental characterization and analysis of the large contributions of ions and water to the thermodynamics of binding IHF to H' DNA. *J Mol Biol* 377:9–27
- Vivas P, Kuznetsov SV, Ansari A (2008) New insights into the transition pathway from nonspecific to specific complex of DNA with *Escherichia coli* integration host factor. *J Phys Chem B* 112:5997–6007
- Wang S, Cosstick R, Gardner JF, Gumpert RI (1995) The specific binding of *Escherichia coli* integration host factor involves both major and minor grooves of DNA. *Biochemistry* 34:13082–13090
- Wedel A, Weiss DS, Popham D, Dröge P, Kustu S (1990) A bacterial enhancer functions to tether a transcriptional activator near a promoter. *Science* 248:486–490
- Weisberg RA, Freundlich M, Friedman D, Gardner J, Goosen N, Nash H, Oppenheim A, Rouvière-Yaniv J (1996) Nomenclature of the genes encoding IHF. *Mol Microbiol* 19:642
- Yang CC, Nash HA (1989) The interaction of *E. coli* IHF protein with its specific binding sites. *Cell* 57:869–880
- Yang SW, Nash HA (1995) Comparison of protein binding to DNA in vivo and in vitro: defining an effective intracellular target. *EMBO J* 14:6292–6300

Chapter 17

Role of HU in Regulation of *gal* Promoters

Dale E.A. Lewis, Sang Jun Lee, and Sankar Adhya

Abstract HU is one of the histone-like DNA binding proteins, which are involved in maintaining the nucleoid structure in bacteria. HU has also been shown to participate in transcriptional regulation of specific promoters. In this chapter, we provide an overview of the mechanism of HU action in the transcriptional regulation of the two promoters of the *gal* operon in *E. coli* by providing results of genetic, biophysical and biochemical experiments both *in vivo* and *in vitro*.

Keywords Anti-parallel DNA Looping • Atomic Force Microscopy • *gal* Operon • GalR • HU

17.1 Introduction

Communication between DNA binding proteins bound at distal sites via DNA looping influences the promoter-RNA polymerase function for regulating gene transcription either positively as in the *gln* operon (Reitzer and Magasanik 1986) negatively as in the *gal* (Irani et al. 1983), *ara* (Dunn and Schleif 1984; Schleif 1987; Lee and Schleif 1989), *deo* (Dandanell and Hammer 1985), and *lac* (Mossing and Record 1986; Kramer et al. 1987, 1988) operons or both positively and negatively as in bacteriophage λ (Griffith et al. 1986; Ptashne 1986). DNA looping frequently needs the aid of an architectural DNA-binding protein, which binds at the apical region of the loop to provide a critical bend and energetic support to the DNA structure. The role of architectural proteins in forming looped DNA structure is widespread. We will discuss the role of an architectural protein in DNA looping using the *gal* operon of *Escherichia coli* as an example. DNA looping in the *gal* system has been extensively studied using molecular genetic, biochemical, and physical methods in order to understand how DNA looping

D.E.A. Lewis, S.J. Lee, and S. Adhya (✉)

Laboratory of Molecular Biology, National Cancer Institute, Bethesda, MD 20892, USA
e-mail: adhyas@mail.nih.gov

influences transcription regulation and to determine the kinetic and thermodynamic parameters of loop formation. In *gal*, the nucleoid protein, HU, provides the architectural role in DNA looping (Aki et al. 1996). The operon specific repressor protein (GalR), HU, and supercoiled *gal* promoter DNA form a higher-order complex structure, termed “repressosome”, containing a DNA loop, which represses the *gal* promoters (Adhya et al. 1998). Although HU is known to regulate repression in other operons (Yang and Larson 1996), in this chapter, we will describe the architectural role of HU in loop assembly and maintenance in the regulation of the *gal* operon.

17.2 The *gal* System

The *gal* operon consists of four structural genes (*galE*, *galT*, *galK* and *galM*), encoding enzymes of the sugar D-galactose metabolism (Fig. 17.1) (Buttin 1963; Echols et al. 1963; Shapiro and Adhya 1969; Bouffard et al. 1994). The genes are transcribed from two overlapping promoters, *P1* and *P2*, separated by 5-bp (Musso et al. 1977; Adhya and Miller 1979). The regulatory region also contains two homologous operators with dyad symmetry. The operators are located 113 bp (11 B-DNA helical turns) from each other on the same face of the helix. The external operator (O_E) is centred at position -60.5 upstream of the promoters, and the internal operator (O_I) at position $+53.5$ within the *galE* gene (Irani et al. 1983; Majumdar and Adhya 1984). GalR is a dimer of 37 kDa and regulates the promoters by binding to the two operators (Buttin 1963; Adhya and Echols 1966; von Wilcken-Bergmann and Muller-Hill 1982; Majumdar and Adhya 1984, 1987).

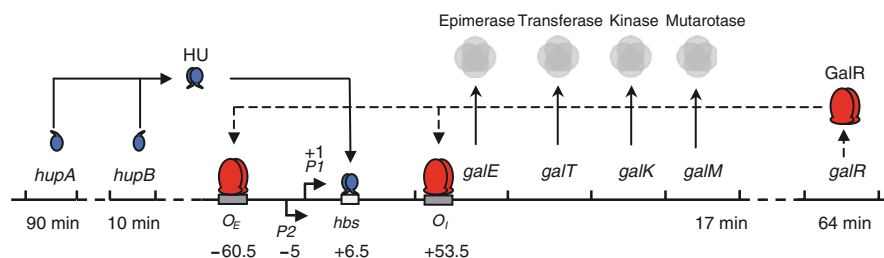


Fig. 17.1 The *gal* regulon. The transcription start point (tsp) of *P1* is $+1$ and that of *P2* is -5 . The nomenclature is relative to the *tsp* of *P1*, with numbers downstream *P1* as positive (+) and those upstream as negative (-). The *gal* regulon is located at 17 min on the *E. coli* genomic map. The *gal* genes (ETKM) are transcribed from the promoters. Their gene products (Epimerase, Transferase, Kinase and Mutarotase) are in gray above each product. HU ($\alpha\beta$) is dimer of gene products from *hupA* (90 min) and *hupB* (10 min), and binds to the histone binding site (*hbs*) centered at position $(+6.5)$. GalR (red) is made from an unlinked gene, *galR* at 64 min, and binds to O_E (-60.5) and O_I ($+53.5$)

Table 17.1 *In vivo* analysis of GalR and LacI effects on the repression of *gal* promoters (Haber and Adhya 1988)

Operators	GalR	LacI
$O_E^G - O_I^G$	R	NR
$O_E^G - O_I^L$	NR	NR
$O_E^L - O_I^G$	NR	NR
$O_E^L - O_I^L$	NR	R

R, Repressible; NR, Non-repressible; O^G , gal operator, O^L , lac operator, O_E , external operator; O_I , internal operator

17.3 In Vivo Evidence for DNA Looping

In vivo, GalR binding to the two operators is required for the repression of the promoters. Conversion of either *gal* operator to a heterologous operator, say *lac* operator, which binds to its cognate repressor, LacI (von Wilcken-Bergmann and Muller-Hill 1982; Kuhnke et al. 1986; Weickert and Adhya 1992) fails to repress the *gal* promoters although both GalR and LacI are present in the cell, suggesting that mere occupation of operator sites is not sufficient for repression (Table 17.1) (Haber and Adhya 1988). However, conversion of both operators to *lac* operators restores normal repression when LacI is present, indicating that repression requires an interaction between operator-bound repressor molecules, which would generate a DNA loop. Interaction occurs only between the same repressor proteins.

17.4 Isolation of a Co-factor for Transcription Repression

An interaction between two proteins while binding to DNA is one of the hallmarks of co-operative binding. Although a GalR-GalR interaction is strongly indicated in vivo, in vitro binding of GalR to O_E and O_I does not show co-operativity at physiological conditions (Brenowitz et al. 1990), and fails to repress both promoters simultaneously (Fig. 17.2a) (Choy and Adhya 1992). GalR binding to O_E represses *P1* and enhances *P2* by inhibiting and promoting open complexes formation respectively (Kuhnke et al. 1986; Choy and Adhya 1992; Choy et al. 1997; Roy et al. 2004). However, simultaneous repression of both promoters requires DNA looping or “repressosome” formation formed by HU, GalR, and supercoiled DNA. In contrast, LacI shows co-operative binding (Brenowitz et al. 1991), transcription repression of both *gal* promoters in vitro in a O_E - O_I DNA molecule in which the *gal* operators were replaced by *lac* operators under similar conditions (Fig. 17.2b) (Choy et al. 1995b). A single LacI molecule binds to the converted *lac* operators simultaneously and efficiently represses the promoters as expected from in vivo results. Does a difference in protein structures account for the difference in repression ability between GalR and LacI in vitro? Is there an additional factor that assists GalR in the repression of the promoters in vivo that is absent in the in vitro conditions?

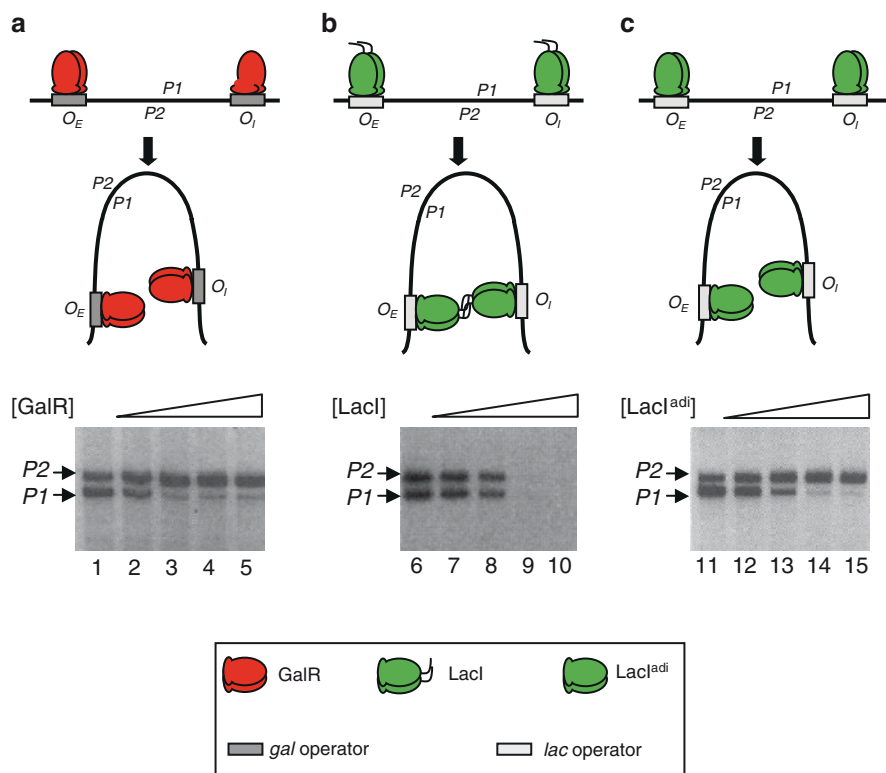


Fig. 17.2 RNAs made by *gal* promoters. (a) Unstable looping model of GalR (red) with gal operator (dark grey rectangle). $P1$ RNAs were repressed while $P2$ was enhanced in the presence of increasing concentration of GalR (lanes 1–5). (b) Stable looping model of LacI (green) and lac operators (light grey rectangle) (lanes 6–10). LacI contains a four-helix bundle as tail. (c) Unstable looping model of LacI^{adi} in the presence of increasing protein concentration (lanes 11–15). LacI^{adi} lacks the four-helix bundle of LacI. (a) and (c) were from Choy et al. (1995a) and (b) from Choy et al. (1995b)

Unlike the dimeric GalR, which binds to each operator separately, LacI is a homotetramer. The C-terminal of LacI additionally contains three leucine-heptads (LAPNTQTASPRALADSLMQLARQVSRLESGQ), which form an anti-parallel four-helix bundle during tetramerization enabling the molecule to bind to two operator sites simultaneously (Alberti et al. 1991, 1993; Friedman et al. 1995; Lewis et al. 1996). The tetramerization helps formation and stabilization of a DNA loop by LacI. The importance of the LacI C-terminus in DNA looping is evident in the behavior of a mutant LacI, LacI^{adi}, which lacks the leucine-heptad sequence and forms very weak tetramers (Oehler et al. 1990). Its binding to O_E and O_I is non-cooperative (Brenowitz et al. 1991). The binding also does not lead to repression, supporting the idea that a stable tetramer of LacI aids DNA loop formation, and hence repression (Fig. 17.2c) (Choy and Adhya 1992; Choy et al. 1995a). GalR lacks the 17-amino acid C-terminal domain of LacI. A DNA-bound GalR dimer-dimer interaction essential for the proposed DNA looping by some domain of the

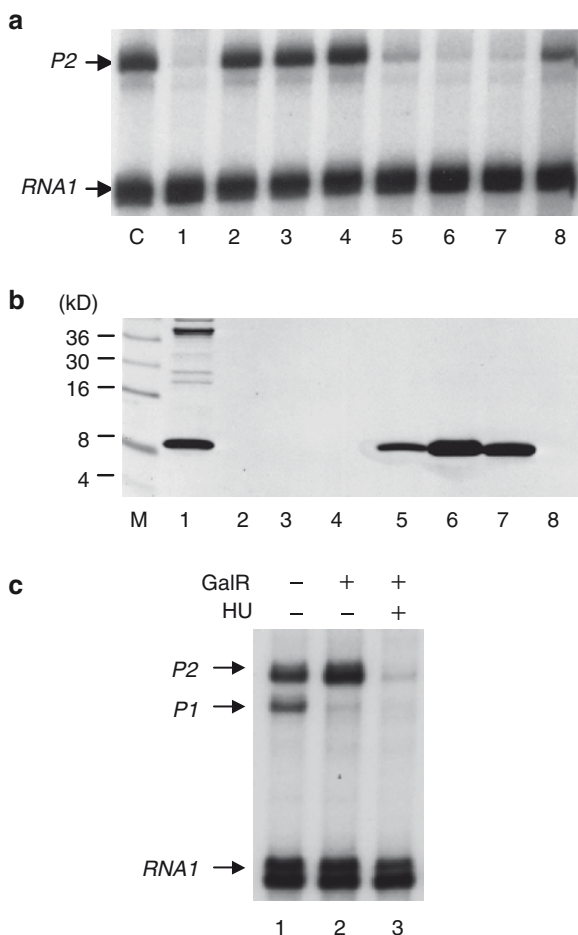


Fig. 17.3 Isolation of a looping factor. **(a)** RNAs were made from *galP2* in the presence of GalR (40 nM) with eight fractions (lanes 1–8) isolated from crude cell extract of *E. coli*. Lane c is control for *P2* transcripts in the absence of GalR. RNAI is an internal control RNA made from the plasmid origin of replication. **(b)** SDS-page of the protein fractions isolated from the crude cell extract above. M is the molecular weight markers. **(c)** RNAs made from *P2* and *P1* in the presence of purified GalR (80 nM) and HU (40 nM). **(a)** and **(b)** were from Aki et al. (1996) and **(c)** from Lewis (2003)

protein may not be stable enough to compensate for the energetic cost for bending and potential twisting of the intervening DNA. If the perceived GalR-GalR tetramer formation is unfavorable, it may need another factor, present in vivo, to stabilize the complex and achieve DNA looping. To identify such a factor, in vitro transcription assays with purified RNA polymerase and GalR were performed on a DNA template containing the *gal P2* promoter in the presence of crude cell extract (S150) to mimic the in vivo conditions (Aki et al. 1996). The presence of crude cell extract indeed shows repression of the *P2* promoter in vitro, as proposed. The crude extract was fractionated to purify the repression co-factor (Fig. 17.3a). The co-factor

migrates as a 9 kDa protein on an SDS gel (Fig. 17.3b; lanes 1, 5–7). Determination of the N-terminal amino acid sequences of the co-factor revealed it to be the nucleoid-associated protein HU (Rouvière-Yaniv 1978). When purified HU is added to in vitro transcription assays, complete repression of both *gal* promoters is observed (Fig. 17.3c). Repression of *P2* occurs only in the presence of both GalR and HU. Other architectural proteins such as IHF, HMG1, FIS and H-NS were not effective in bringing out GalR-mediated repression in *gal* suggesting that HU is specific for *gal* DNA loop (Aki and Adhya 1997) (Lewis, D.E.A. unpublished result). In addition to HU, DNA looping by GalR also needs a supercoiled DNA template both in vivo and in vitro (Aki et al. 1996; Lewis et al. 1999). Linear DNA does not show repression in the presence of GalR and HU (Aki et al. 1996). As expected, repression requires GalR binding to both operators. Deletion of either O_E or O_I abolishes repression in vitro (Lewis 2003).

17.5 Properties of HU

HU is a heterodimer of HU- α (encoded by the *hupA* gene) and HU- β (encoded by the *hupB* gene) proteins and present in about 30,000 copies per cell (Rouvière-Yaniv and Gros 1975). The α and β subunits show 70% identity in amino acid sequences (Fig. 17.4) (Rouvière-Yaniv and Gros 1975; Kano et al. 1986, 1987; Drlica and Rouvière-Yaniv 1987). HU binds non-specifically to DNA in micromolar concentrations but recognizes aberrant DNA structures with higher affinities (Bonnefoy and Rouvière-Yaniv 1991; Pontiggia et al. 1993; Tanaka et al. 1993; Bonnefoy et al. 1994; Castaing et al. 1995). Although purified as a factor in transcriptional regulation in *gal*, the protein was already known to be a component of the bacterial nucleoid and to be involved in DNA replication, transposition, recombination, repair and cell division (Craigie and Mizuuchi 1985; Surette et al. 1987; Pettijohn 1988; Lavoie and Chaconas 1993).

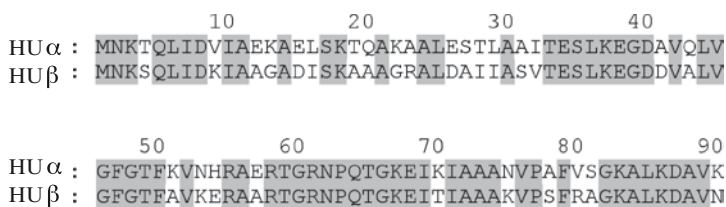


Fig. 17.4 Alignment of HU α and HU β protein sequences. Identical amino acids found in both proteins are shaded in grey

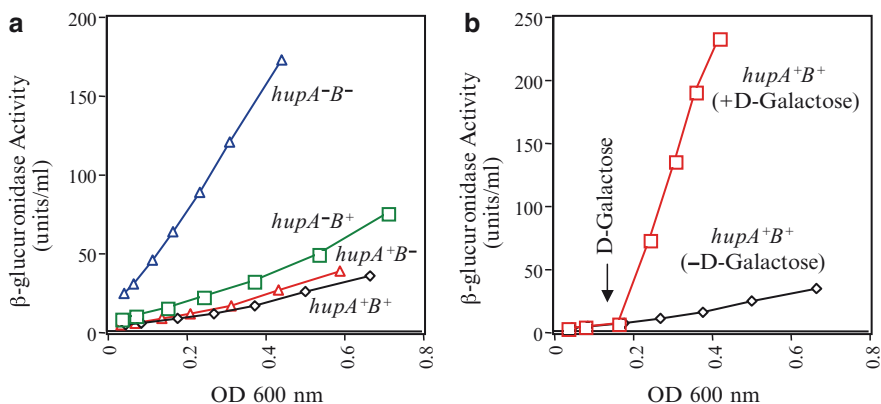


Fig. 17.5 Effect of *hup* deletions on *galP2-gusA* activity. (a) Differential rate of β-glucuronidase synthesis in wild-type (*hupA⁺B⁺*), and mutants (*hupA⁺B⁻*), (*hupA⁻B⁺*), and (*hupA⁻B⁻*) strains. Cells were grown to various absorbances at 600 nm and the enzyme activities were measured at 405 nm. (b) Wild-type strain in the presence and absence of D-galactose, an inducer of GalR. The arrow indicates the point at which D-galactose was added to the growing culture (Lewis et al. 1999)

17.6 Confirmation of the Requirement of HU for *gal* Repression In Vivo

The role of HU in *gal* repression was investigated in vivo by assaying repression of a reporter gene, *gus* (encoding the enzyme β-glucuronidase) fused to the *galP2* promoter in strains deleted for either *hupA* or *hupB* or both genes (Lewis et al. 1999). While the wild type strain is repressed for *P2*, strains carrying deletion of both *hupA* and *hupB* genes derepressed the promoter fully as did the addition of D-galactose, an inducer of the operon (Fig. 17.5a and b). However, deletion of either *hupA* or *hupB* shows a high level of *P2* repression, suggesting that in addition to the HUαβ heterodimer, the HUαα and HUββ homodimers are quite effective in *P2* repression in vivo.

17.7 HU Binds to *gal* DNA

As mentioned before, HU is mostly a non-specific DNA binding protein. However, by KMnO₄ cleavage assay (Lavoie and Chaconas 1993; Lavoie et al. 1996), HU shows specific binding to *gal* DNA (Aki and Adhya 1997). In this method HU is converted to a chemical nuclease by linking it to EPD-Fe (+++); addition of KMnO₄ to modified HU generates hydroxyl radicals, which cleave DNA within 10 Å of the EPD ligands. HU binding to *gal* DNA needed (i) the presence of GalR, (ii) a DNA carrying both operators, and (iii) supercoiled DNA. HU binding was centered at position +6.5 in *gal* DNA (*hbs*), which is somewhat closer to *O_I* than the midpoint of the operators as indicated in Fig. 17.1. The region is AT-rich, hence it is easily bendable

and perhaps forms a kink or distorted structure, which is recognized by HU. It is noteworthy that whereas the affinity of HU for nonspecific DNA binding is micromolar, for *gal* DNA in the presence of GalR it is nanomolar (Aki and Adhya 1997).

17.8 Demonstration of DNA Looping In Vitro: Helical Phasing of the Operators

The formation of a stable DNA loop by interaction between two DNA-bound proteins may depend on a variety of conditions, such as the proper angular orientation of the protein binding sites on DNA to facilitate protein-protein interactions (Hochschild and Ptashne 1986; Kramer et al. 1987; Wang and Giaever 1988; Lee and Schleif 1989; Law et al. 1993), the size of the loop (Muller et al. 1996), and the superhelicity of the DNA (Borowiec et al. 1987; Kramer et al. 1988; Lobell and Schleif 1990). Because shorter DNA resists torsional change (Shore and Baldwin 1983), the proper angular orientation of the two binding sites is more important with relatively short loop size. It has been suggested that a loop size of less than 150 bp strictly depends upon proper phasing of the protein binding sites, less so for 200 bp, and not at all for 400 bp and longer (Virnik et al. 2003). The precise size of the helix repeat also depends to a certain extent on the DNA sequence of the loop (Muller et al. 1996). Whether HU-mediated transcription repression achieved in *gal* by GalR is because of DNA loop formation was tested by measuring the efficiency of *P*₂ repression in DNA molecules with changing relative angular orientations of the two binding sites. *O*_E and *O*_I are 113-bp apart (11 B-DNA helical turns) and are located on the same face of the DNA, which is ideal for DNA looping. Engineered DNA templates containing base pairs addition (1–21) between *hbs* and *O*_I or deletion (2–12) between *O*_E and *hbs* were tested for in vitro transcription repression by GalR and HU (Fig. 17.6a). The results clearly showed that repression efficiency changes with an increasing number of base-pair additions or deletions in a periodic fashion commensurate with DNA helical repeat. When both operators are in phase (103, 113, 123 and 134-bp apart), repression of *gal* is observed; when they are out of phase, repression is lifted either partially or completely (Fig. 17.6a) (Lewis and Adhya 2002). These experimental results strongly suggest that DNA looping in *gal* and simultaneous repression of both promoters in *gal* needs DNA looping. Even changing the distance between the two GalR binding sites by one base-pair deletion or addition decreases repression efficiency; in order for HU to stabilize the GalR-GalR dimers at *O*_E and *O*_I, both operators have to be perfectly aligned on the same face of the DNA. Furthermore, insertion of one (10 bp), two (21 bp), four (42 bp), five (52 bp), six (63 bp), seven (73 bp), eight (84 bp), nine (94 bp) or ten (105 bp) extra helical turns to the native intra-operator distance of 11 helical turns between *O*_E and *O*_I shows that repression was tight up to the addition of eight or more extra helical turns (to a total intra-operator distance of 19 helical turns) but decreases with a further increase in the number of helical repeats (Fig. 17.6b) (Virnik et al. 2003). The failure to efficiently repress in the latter cases is perhaps because of reduced DNA loop formation even in the presence of HU.

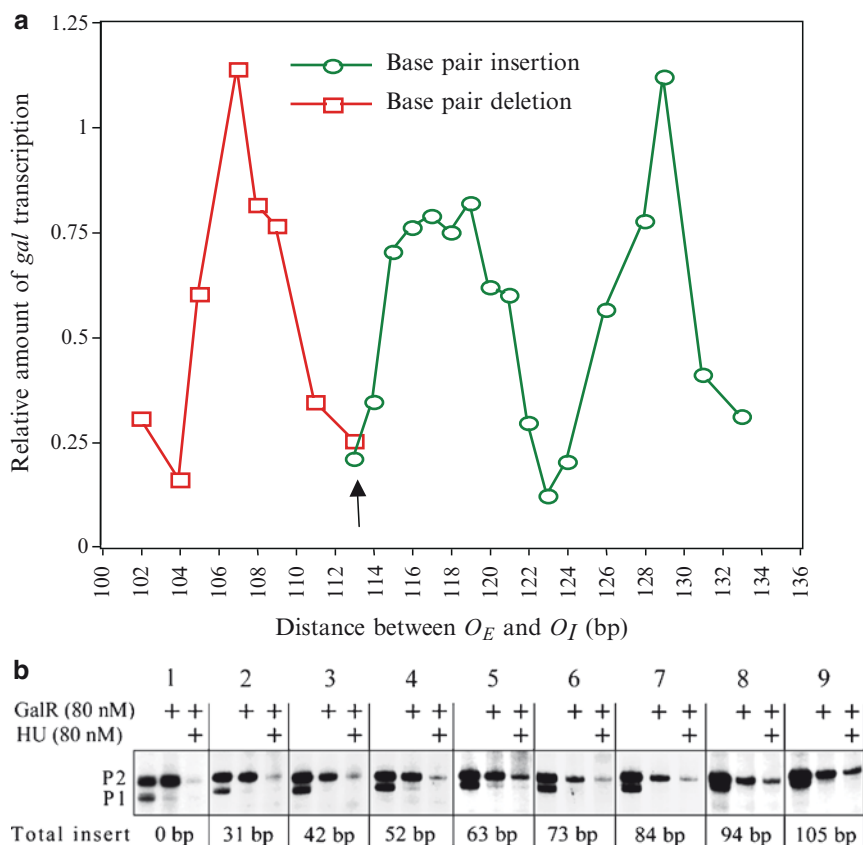


Fig. 17.6 Helical relationship of *gal* repression and operators distance. **(a)** A graph showing the correlation between P_2 repression and an operator distance of 102–132 bp between O_E – O_I ; the wild type distance between O_E – O_I is 113-bp, indicated by the arrow. Deleted base pairs are in red (*squares*) and inserted base pairs are in green (*circles*). The relative amount of transcripts is an indication of promoters' repression (a lower value indicates more repression, while a high value indicates less or no repression). **(b)** RNAs made from *gal* promoters in the presence of GalR and HU as the distance between O_E – O_I increases by an additional 10 more helical turns, assuming 10.5 bp per helical turn. **(a)** From Lewis and Adhya (2002) and **(b)** from Virnik et al. (2003)

Incidentally, GalR/HU mediated DNA looping as a periodic function of the DNA length between the two GalR binding sites allowed estimation of the size of a DNA helical turn under in vitro conditions to be ~10 bp. A helix repeat of 10.5 bp has also been found in DNA looping by LacI and bacteriophage λ cI repressors in vitro (Hochschild and Ptashne 1986; Kramer et al. 1987). The in vitro results of DNA helix repeat size, based upon protein-mediated DNA looping, are consistent with the estimation of the helix size of about 10.5 bp by other in vitro assay systems using both circular and linear DNA (Wang 1979; Rhodes and Klug 1980; Tullius and Dombroski 1985; Hahn et al. 1986). Dependence of loop formation on the relative angular orientation of the two protein binding sites in a periodic fashion and thereby estimation of the size of the helix repeat has also been studied in several

systems in vivo including an artificial construct in which a heterologous promoter was spanned by *gal* operators, O_E and O_I (Kramer et al. 1988; Lee and Schleif 1989; Law et al. 1993; Muller et al. 1998; Perez et al. 2000). The size of a DNA helical turn by in vivo estimation was shown to be within 11.0–11.3 bp. Thus, there is a discrepancy of about one base pair in helix repeat size of DNA structure in vitro and in vivo. The reason for this difference needs further investigation.

17.9 GalR Dimer–Dimer Interface Involved in DNA Looping

Although GalR exists as a dimer in solution, there are several lines of evidence demonstrating that GalR dimers interact and form tetramers. First, a structure-based genetic study of the interaction of the two operator-bound GalR dimers identified several amino acid residues, which are involved in dimer–dimer interaction and thus define the GalR tetramerization interface (Geanacopoulos et al. 2001). Second, a GalR mutant (R282L) has been isolated and characterized in which one of the amino acid residues (R282) identified to participate in dimer–dimer contacts represses transcription in the absence of HU both in vivo and in vitro (Semsey et al. 2002). The repression involves the same DNA looping; deletion of either operator makes the substituted GalR ineffective in repression. The R282L substitution increases normal affinity between two GalR dimers, allowing looping without HU. This suggests that GalR dimers can directly interact; HU (and DNA supercoiling) stabilize the GalR–GalR interaction. Third, after treatment of GalR alone with a chemical homobifunctional imidoester cross-linking agent, dimethyl suberimidate, followed by SDS-gel electrophoresis, the protein shows the existence of monomers, dimers, trimers and tetramers clearly demonstrating the ability of GalR to tetramerize (Semsey et al. 2002). Fourth, Analytical ultracentrifugation studies demonstrated the ability of GalR to form tetramers (Roy et al. 2005).

17.10 How DNA Looping Brings About Repression

As described above, a closed DNA loop represses the promoters located within the loop. It is suggested that the DNA in a topologically closed loop is inflexible to torsional changes and makes the promoters inadequate for transcription by not permitting DNA unwinding needed for open complex formation (Choy et al. 1995a).

17.11 Mechanism of Action of HU

If HU is a co-factor to bring about *gal* DNA looping, what is the role of HU? One can envision three ways by which HU can help *gal* DNA looping (Fig. 17.7): (a) HU bends DNA to facilitate looping; (b) HU acts as a mediator of GalR–GalR interaction in DNA looping; (c) HU bends DNA as well as mediates GalR–GalR interaction. In (a), the interaction energy of tetramer formation by two DNA bound GalR

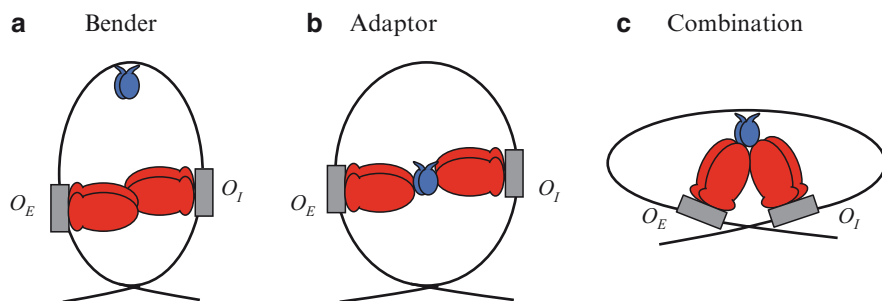


Fig. 17.7 Model of HU role in *gal* DNA looping. (a) HU acting as a bender to bring about GalR-GalR tetramerization, resulting in looping repression of *gal* promoters. (b) HU acting as an adaptor to achieve GalR tetramerization. (c) HU acting both as a bender and adaptor during loop formation. The symbols are described in the legend of Fig. 17.1

dimers is too low to overcome the energy required for DNA bending and possibly twisting needed for DNA loop formation. Although DNA supercoiling overcomes part of the energetic barrier, HU can further help by binding at an architecturally critical position in the DNA to bend the DNA and sustain the loop. HU is known to be a DNA bending protein (Drlica and Rouvière-Yaniv 1987; Pettijohn 1988; Bonnefoy and Rouvière-Yaniv 1991). In (b), GalR dimers cannot form tetramers by themselves thus necessitating the need for a mediator to bring them together. HU may play that role. However, the demonstration of GalR tetramer formation by ultracentrifugation studies (Roy et al. 2005), the ability of mutant GalR dimers to interact with each other (Geanacopoulos et al. 1999), and the demonstration of a specific HU-DNA interaction in the GalR loop rule out a simple adaptor model (Aki et al. 1996; Aki and Adhya 1997). In (c), HU plays the role of both DNA bender and GalR adaptor. Given the limited size of the DNA segment (113 bp), a loop containing three proteins bound to DNA at spatially separated sites and simultaneously contacting each other would be energetically and/or sterically prohibitive (Majumdar et al. 1987). We will distinguish between models (a) and (c) below.

17.12 Assembly of the *gal* Loop: Role of HU

Models (a) and (c) above are distinguished by the prediction that in the latter, there would be an interaction between GalR and HU. In fact, studies of HU binding to DNA revealed that there is a tripartite co-operativity between GalR and HU in binding to *gal* DNA; binding of HU to *gal* DNA is absolutely dependent upon binding of GalR to both *gal* operators, and HU binding in turn results in increasing the strength of GalR binding (Aki and Adhya 1997). The binding of GalR to the operators in the absence of HU is non-cooperative (Brenowitz et al. 1990). The basis of a synergistic binding of GalR and HU could be an interaction between GalR and HU. Co-immuno-precipitation of GalR and HU, indeed, shows a specific interaction between the two proteins both in crude extracts and with purified proteins that was

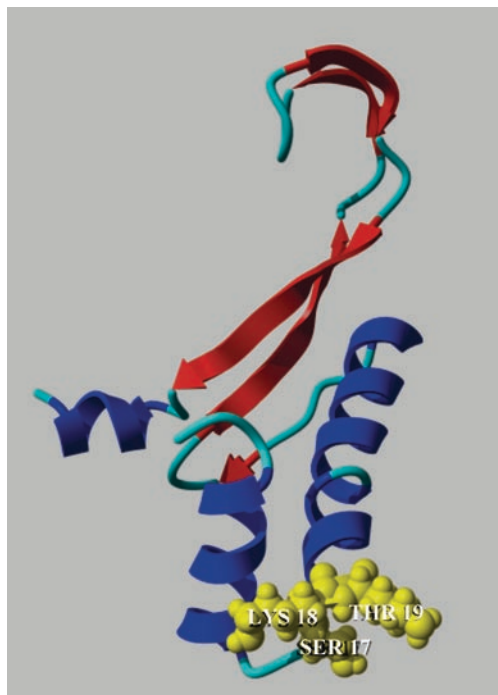


Fig. 17.8 Model of HU α showing the locations of Serine17, Lysine18, and Threonine19 (Kar and Adhya 2001)

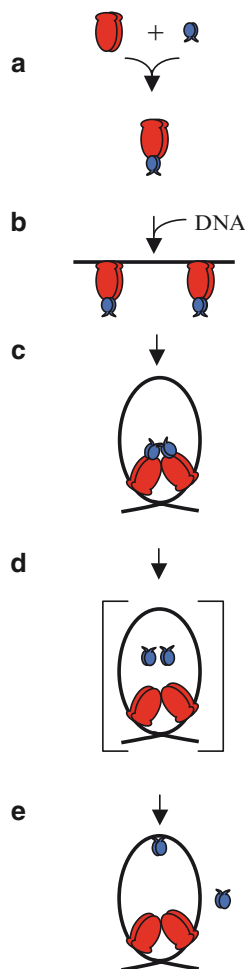
not evident between an HU homolog, IHF (Oberto et al. 1994), and GalR (Kar and Adhya 2001). The amino acid residues in HU that define the contact(s) between HU and GalR have been identified by characterizing mutants of HU α (S17P, K18A, and T19D) that show inefficient co-immunoprecipitation with GalR and do not help *gal* repression efficiently both in vitro and in vivo (Fig. 17.8) (Kar and Adhya 2001). Electrophoretic mobility-shift assays to test the DNA-binding properties of the HU mutants show that the mutants are proficient in DNA-binding.

The crystal structure of the *B. stearothermophilus* HU homodimer (HBs), which shares a 59% sequence homology with *E. coli* HU heterodimer (Drlica and Rouvière-Yaniv 1987), was used to model the *E. coli* structure (Kar and Adhya 2001). The amino acid residues S17, K18, and T19 in *E. coli* HU α are located contiguously in a small turn between the first and second alpha helices (Fig. 17.7). This region lies on the opposite face of the DNA-binding surface, in a prominently accessible portion of HU. All three amino acids have solvent-exposed side chains and are likely candidates to contact GalR. No information is available related to the corresponding residues in GalR.

Given the facts discussed above, a pathway for HU-mediated DNA looping in *gal* emerges (Fig. 17.9) (Kar and Adhya 2001). (a) GalR dimers recruit HU. (b) GalR–HU complexes bind to the operators. (Step b could precede step a.) (c) The

Fig. 17.9 A pathway of GalR and HU to DNA looping in the presence of supercoiled DNA. Steps (a)–(e) are described in the text. The symbols are described in the legend of Fig. 17.1

Pathway to DNA looping



operator-bound GalR–HU complexes transiently interact to form tetramers via GalR dimer–dimer interaction and generate a DNA loop that has a distortion around the apical region (*hbs*). (d) Spontaneous dissociation of HU from GalR increases the local concentration of HU. (e) Because HU has a stronger affinity to distorted DNA (~ 10 nM) than to GalR (~ 10 μ M), the dissociated HU preferentially binds to the architecturally critical but transiently bent/distorted DNA, stabilizing the loop. By this scenario, HU binds and stabilizes the loop by binding to an architecturally critical position in DNA. This pathway is consistent with the energetics of the various interactions reported above. The concentration of HU in the cell is not high enough to interact with a rare and transient target in *gal* DNA. The recruitment of HU by GalR as a piggyback in *gal* looping to increase local concentration for DNA binding may be

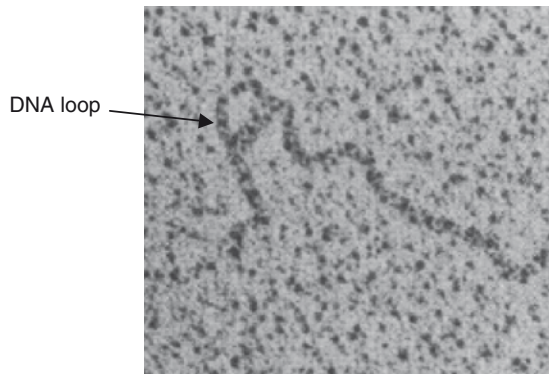


Fig. 17.10 Electron microscopy image of DNA looping with LacI and *lac* operators on a linear DNA (reproduced with permission from Cold Spring Harbor Laboratory Press (Mandal et al. 1990))

an example of a common theme in the formation of similar higher order structures involving both architectural and sequence-specific DNA binding proteins in other DNA transaction reactions (Charlier et al. 1980; Kabsch et al. 1982; Borowiec et al. 1987; Lewis et al. 1996; Muller et al. 1996; Dickerson 1998).

17.13 Electron Microscopic Evidence of a DNA Loop

When electron microscopy is used to study loop formation on a *gal* DNA fragment containing either the $O_E^G - O_I^L$, $O_E^L - O_I^G$ or $O_E^L - O_I^L$ genotype in the presence of GalR (in the absence of HU), LacI^{adi} or LacI, DNA loops of the expected size are observed only with the $O_E^L - O_I^L$ DNA and wild type LacI tetramer protein but not with LacI^{adi}, a mutant dimer protein (Fig. 17.10) (Mandal et al. 1990). Looping is never observed with the $O_E^L - O_I^L$ DNA in the presence of GalR. This result was expected because LacI tetramer represses transcription in vitro but LacI^{adi} and GalR do not (Choy and Adhya 1992; Choy et al. 1995a).

17.14 Evidence for an Anti-parallel *gal* DNA Loop: Modeling

We mentioned above that two GalR dimers form a V-shaped tetramer stabilized by stacking interactions. DNA loop formation by GalR tetramer as predicted by genetic data mentioned above can align the operators in four possible arrangements (Fig. 17.11a and b), two antiparallel loops (A1 and A2) and two parallel loops (P1 and P2) (Geanacopoulos et al. 2001; Virnik et al. 2003; Semsey et al. 2004; Lia et al. 2008). Antiparallel loops mean that the 5' end of one operator is arranged in the opposite direction of the 5' end of the other operator (A1 and A2). In parallel loops, the 5' end of one operator is arranged in the same direction as the 5' end of the other operator (P1 and P2).

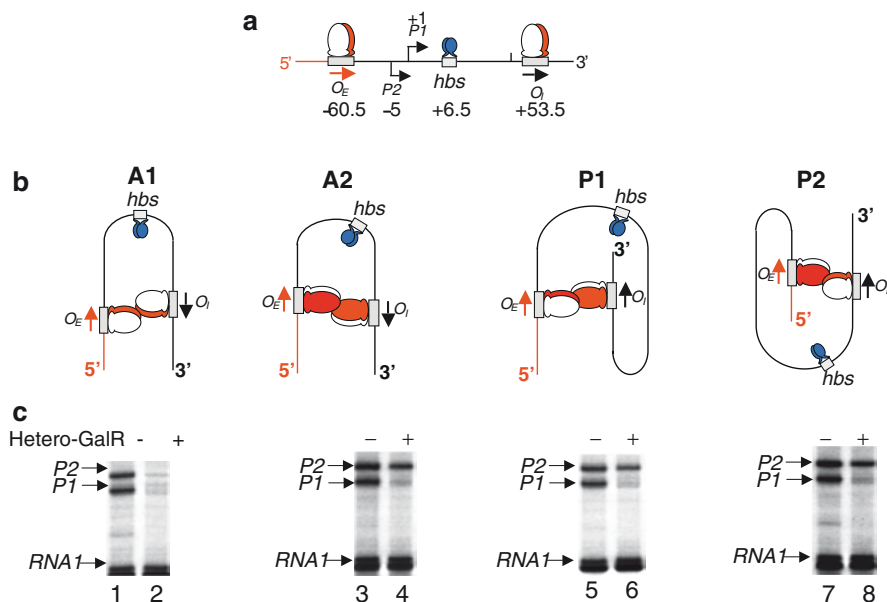


Fig. 17.11 Models of DNA loops formed by GalR. (a) Regulatory region of the *gal* regulon as described above in Fig. 17.2. The monomer of GalR is red and the other monomer is white. The red and black arrows indicate the direction from 5' to 3' of the operators, O_E and O_I , respectively. (b) Possible loops of GalR-DNA interactions. A1 and A2 represent two forms of “antiparallel” loops and P1 and P2 represent two forms of “parallel” loops. (c) RNAs made from *galP1* and *P2* with GalR heterodimer and different operators to generate the four possible loops: A1 (lanes 1–2), A2 (lanes 3–4), P1 (lanes 5–6) and P2 (lanes 7–8) (reproduced with permission from Cold Spring Harbor Laboratory Press (Semsey et al. 2004))

These four possibilities were analyzed by stereochemical model building using an empirical elastic model of DNA (Geanacopoulos et al. 2001). First, the GalR dimer structure was modeled after the x-ray structure of a GalR homolog, PurR (Schumacher et al. 1994). The model of the HU-DNA complex was based on the crystal structure of the IHF-DNA complex (Rice et al. 1996). GalR tetramer configurations were constructed using the sequence-dependent structural parameters of the interoperator DNA and conformation changes caused by GalR and asymmetric HU binding. Evaluation of their DNA elastic energies gives unambiguous preference to a loop structure in which the two *gal* operators adopt an antiparallel orientation. The antiparallel modes (A1 and A2) of DNA looping proved to be more favorable energetically than the parallel modes (P1 and P2) (Table 17.2). From an energetic point of view, formation of the optimized parallel loop, P2, results in a five-fold increase in the bending energy compared to the optimized antiparallel loop, A1.

The A1 loop, unlike the parallel loops, would unwind the inter-operator DNA and would be stabilized by negative supercoiling, an absolute requirement for repressosome formation (Aki and Adhya 1997). The A1 structure is consistent with the asymmetric (closer to O_E than to O_I) binding of HU when the 5' end of the loop is longer than

Table 17.2 Energy of 113 bp DNA looping for the optimized GalR loops (Geanacopoulos et al. 2001)

Loop ¹	$-\Delta G_{\text{total}}^2$
A1	9.4
A2	25.5
P1	92.3
P2	25.5

¹The four loops are schematically presented in Fig. 11.

²Energy values are in kcal mol⁻¹. $-\Delta G_{\text{total}} = -\Delta G_{\text{bend}} - \Delta G_{\text{twist}}$

the 3' end. The sufficiency of negative supercoiling to stimulate binding of an HU heterodimer to a specific site (Kobryn et al. 1999) suggests that the unwinding of the interoperator DNA in the A1 loop is important for HU binding to its site, *hbs*.

17.15 In Vivo Evidence for an Anti-parallel Loop

To distinguish which loop arrangement is the correct structure in the repressosome, heterodimers of GalR were constructed and the operators designed to orient GalR in one fixed direction and thus generate the four possible loop structures (Semsey et al. 2004). The results showed that simultaneous repression of *gal P1* and *P2* operators by the GalR heterodimer is achieved only when the DNA trajectory was in A1 form (Fig. 17.11c, lanes 1–2) and not in the A2 form (lanes 3–4) or in the two parallel forms (P1 and P2) (lanes 5–8), confirming the energetic prediction that A1 loop formation requires least energy. Moreover, HU binds to its site located at +6.5 in the A1 model. In the A2 model, HU would bind to a site that is not located at the apex of the loop and would not be able to stabilize the loop in the A2 case. This result shows that the antiparallel form is the correct trajectory of GalR in DNA looping. However, as discussed below, the energetic calculations for loop formation by single DNA molecule analysis indicates that both A1 and A2 loops form with similar energies. The discrepancy is discussed later.

17.16 Evidence of Anti-parallel Loops: AFM Studies

The distinction between the two trajectories that are most feasible from energetic considerations (A1 and P2) was studied by AFM which confirmed the anti-parallel nature of the *gal* loop (Fig. 17.12) (Virnik et al. 2003). For convenience of comparison of theoretical prediction and experimental results, the two models (a) and (c) are presented as parts incorporated into imaginary DNA minicircles depicted with broken lines. Arrows are used to indicate the 5-prime to 3-prime directions for the operators, which concur with the directions of transcription from the *gal* promoters. It is clear from the AFM images that all the loops observed in the presence of GalR and HU conform to a trajectory of an anti-parallel A1 loop (b); loops as expected for a parallel trajectory were never observed, giving credence to the idea that anti-parallel loops are more stable.

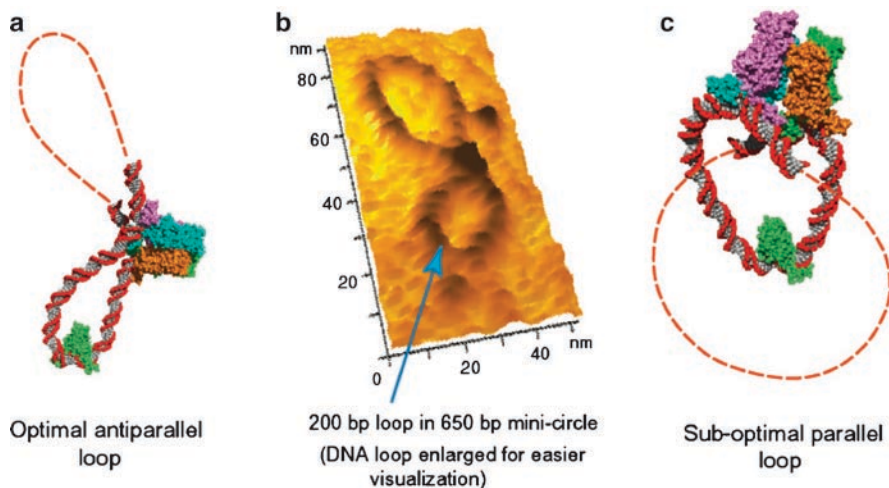


Fig. 17.12 Atomic force microscopy of GalR-HU loop with minicircle. (a) Proposed model of “antiparallel” loop. (b) Experimental AFM image of “antiparallel” loop. (c) Proposed model of “parallel” loop. The arrows indicate the direction of the operators from 5' to 3'. The blue and red arrows represent O_I and O_E , respectively (Virnik et al. 2003)

17.17 Stability of the *gal* Loop: Single DNA Molecule Analysis

By using single-molecule micromanipulation to generate and finely tune tension in DNA molecules, the kinetics, thermodynamics, and supercoiling dependence of GalR/HU-mediated DNA looping have been characterized. The factors required for *gal* DNA looping in single-molecule experiments (HU, GalR and DNA supercoiling) correspond exactly to those essential for *gal* repression in vitro and in vivo. Magnetic tweezers were used to determine the thermodynamics of loop formation on DNA tethered at one end (Fig. 17.13) (Lia et al. 2003, 2008). The technique involves detecting the transition between unlooped (L) and looped (L-1) structures, where L represents the length of the tethered particle and L-1, represents the change in the length of the particle due to looping. By using an optimal force (F) of 0.88 pN (to keep the DNA somewhat stretched), and superhelical density (σ) of -0.03 (to keep the DNA slightly untwisted), looping and unlooping is observed (Lia et al. 2003). DNA molecules that are negatively supercoiled by at least 3% ($\sigma = -0.03$) and stretched with a force (F) of 0.88 pN intermittently switched between two conformations in the presence of both GalR and HU. No length-changes are observed in the absence of GalR and/or HU or in the presence of D(+)-galactose. Similarly, GalR and HU do not generate looping in molecules containing only one operator, O_E or O_I . The extrapolated mean lifetimes of looped and unlooped conformation in the presence of GalR and HU were <1 ms and ~21 sec, respectively (Lia et al. 2003). Looping and unlooping are not observed when the DNA was relaxed ($\sigma = 0$) or contained positive supercoiled ($\sigma = +0.03$). The associated free energy change involved in the GalR/HU-mediated loop is $\sim -12 k_B T$ or -7.1 kcal/mol (Geanacopoulos

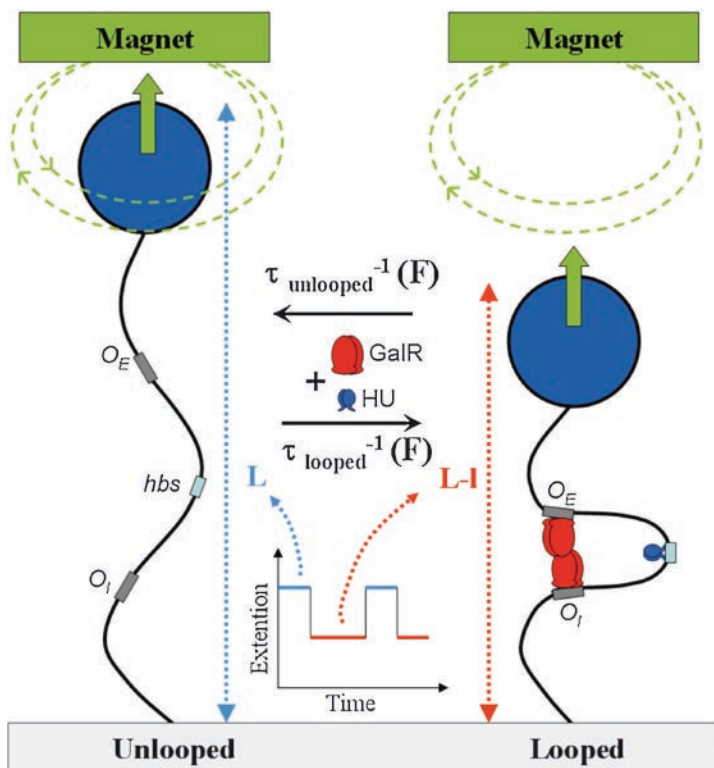


Fig. 17.13 Experimental layout to measure DNA looping by magnetic tweezers. A *gal* DNA molecule is attached to a cover slide and a magnetic force (green arrow) is applied to it. The length of the tethered molecule (*unloop*) is indicated by (L) and that for the loop molecule by (L-1). The transitions from the unloop state (broken blue arrow) to the loop state (broken red arrow) are shown in the middle of the figure

et al. 2001; Lia et al. 2003), which compares to -9.4 Kcal/mole estimated from the modeling studies described above (Geanacopoulos et al. 2001).

17.18 Energetics of A1 and A2 Loops Estimated by Single Molecule Studies

The single DNA molecule assay employing magnetic tweezers when applied to distinguish between locked A1 and A2 arrangements indicate that the A1 and A2 loops formed with similar energies in DNA. The two loops show nearly equivalent probabilities of formation and the change in free energy for loop formation increases with increasing force i.e., tension destabilizes the loops. Therefore, it appears that there is no thermodynamic reason accounting for the observed functional difference in transcriptional repression between A1 and A2 loops. However, the loop lifetimes

determined by extrapolation to zero force may not be relevant. There are suggestions that DNA is under tension in vivo and several motor enzymes, such as RNA polymerase, have been reported to exert large forces on the topologically constrained DNA (Wang et al. 1998; Pennington 2006). Furthermore, DNA molecules more negatively supercoiled (e.g., 6% or $\sigma = -0.06$) have a built-in entropic tension of about 0.5 pN (Charvin et al. 2004).

17.19 “Antiparallel” Loops and Bacterial Nucleoid Structure

A basic concept of DNA loops with “antiparallel” trajectory, which is involved in gene regulation, may help understand structural organization of the bacterial nucleoid. The problem of bacterial genomic DNA compaction has received attention for decades without a solution (Drlica 1987; Pettijohn 1996). There appear to be several factors that influence the nucleoid compaction, including supercoiling of the DNA, macromolecular crowding, RNAs, cellular polyamines, and nucleoid-associated proteins (Murphy and Zimmerman 2000). DNA supercoiling is the initial step in bacterial nucleoid organizational hierarchy. It is very likely that DNA looping is a cohort in the process. Generally speaking, both the “antiparallel” (A1) looping modes used in the *gal* system and the “parallel” (P2) looping modes present in eukaryotic nucleosome may be involved in the DNA compaction process in bacterial nucleoid.

17.20 Summary

The two promoters of the *gal* operon are simultaneously repressed only by looping the DNA containing the promoters; looping makes the promoters inadequate for open complex formation in transcription initiation. Results of transcription repression of the *gal* promoters, obtained by in vitro transcription assays in DNA with enlarged inter-operator distance, show that the HU protein is not essential for loop formation although it clearly facilitates the process. GalR dimers could interact directly and not through an adaptor. In other words, the presence of HU is critical apparently not for closure of the DNA loop by acting as an adapter but for the stabilization of the latter; after its recruitment by GalR, HU stabilizes the architecture of the DNA loop geometry. HU does not affect RNA polymerase activity in the *gal* promoters as a transcription factor. HU shows tripartite-binding cooperativity with GalR dimer binding to the operators in DNA looping (Grove et al. 1996). Probably, increasing the distance between O_E and O_I reduces the resistance to torsional and bending distortions, and shifts the energy balance in favor of a more stable GalR tetramer.

Acknowledgements Much of the work described in this chapter is from the authors' laboratory. We gratefully acknowledge the contributions of many of our past and present colleagues and collaborators too numerous to list. We are supported by the Intramural Research Program of the National Institutes of Health, the National Cancer Institute, the Center for Cancer Research. We thank Dr. Dhruba Chatteraj for suggestions and critical reading of the manuscript.

References

- Adhya S, Echols H (1966) Glucose effect and the galactose enzymes of *Escherichia coli*: correlation between glucose inhibition of induction and inducer transport. *J Bacteriol* 92:601–608
- Adhya S, Miller W (1979) Modulation of the two promoters of the galactose operon of *Escherichia coli*. *Nature* 279:492–494
- Adhya S, Geanakopoulou M, Lewis DE, Roy S, Aki T (1998) Transcription regulation by repressor and by RNA polymerase contact. *Cold Spring Harb Symp Quant Biol* 63:1–9
- Aki T, Adhya S (1997) Repressor induced site-specific binding of HU for transcriptional regulation. *EMBO J* 16:3666–3674
- Aki T, Choy HE, Adhya S (1996) Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of *gal* transcription by GAL repressor. *Genes Cells* 1:179–188
- Alberti S, Oehler S, von Wilcken-Bergmann B, Kramer H, Muller-Hill B (1991) Dimer-to-tetramer assembly of Lac repressor involves a leucine heptad repeat. *New Biol* 3:57–62
- Alberti S, Oehler S, von Wilcken-Bergmann B, Muller-Hill B (1993) Genetic analysis of the leucine heptad repeats of Lac repressor: evidence for a 4-helical bundle. *EMBO J* 12:3227–3236
- Bonnefoy E, Rouvière-Yaniv J (1991) HU and IHF, two homologous histone-like proteins of *Escherichia coli*, form different protein-DNA complexes with short DNA fragments. *EMBO J* 10:687–696
- Bonnefoy E, Takahashi M, Yaniv JR (1994) DNA-binding parameters of the HU protein of *Escherichia coli* to cruciform DNA. *J Mol Biol* 242:116–129
- Borowiec JA, Zhang L, Sasse-Dwight S, Gralla JD (1987) DNA supercoiling promotes formation of a bent repression loop in *lac* DNA. *J Mol Biol* 196:101–111
- Bouffard GG, Rudd KE, Adhya SL (1994) Dependence of lactose metabolism upon mutarotase encoded in the *gal* operon in *Escherichia coli*. *J Mol Biol* 244:269–278
- Brenowitz M, Jamison E, Majumdar A, Adhya S (1990) Interaction of the *Escherichia coli* Gal repressor protein with its DNA operators *in vitro*. *Biochemistry* 29:3374–3383
- Brenowitz M, Mandal N, Pickar A, Jamison E, Adhya S (1991) DNA-binding properties of a *lac* repressor mutant incapable of forming tetramers. *J Biol Chem* 266:1281–1288
- Buttin G (1963) Regulatory mechanisms in the biosynthesis of the enzymes of galactose metabolism in *Escherichia coli* K 12. II. the genetic determinism of the regulation. *J Mol Biol* 7:183–205
- Castaing B, Zelwer C, Laval J, Boiteux S (1995) HU protein of *Escherichia coli* binds specifically to DNA that contains single-strand breaks or gaps. *J Biol Chem* 270:10291–10296
- Charlier M, Maurizot JC, Zaccai G (1980) Neutron scattering studies of *lac* repressor. *Nature* 286:423–425
- Charvin G, Allemand JF, Strick TR, Bensimon D, Croquette V (2004) Twisting DNA: single molecule studies. *Contemp Phys* 45:383–403
- Choy HE, Adhya S (1992) Control of *gal* transcription through DNA looping: inhibition of the initial transcribing complex. *Proc Natl Acad Sci USA* 89:11264–11268
- Choy HE, Park SW, Aki T, Parrack P, Fujita N, Ishihama A, Adhya S (1995a) Repression and activation of transcription by Gal and Lac repressors: involvement of alpha subunit of RNA polymerase. *EMBO J* 14:4523–4529
- Choy HE, Park SW, Parrack P, Adhya S (1995b) Transcription regulation by inflexibility of promoter DNA in a looped complex. *Proc Natl Acad Sci USA* 92:7327–7331
- Choy HE, Hanger RR, Aki T, Mahoney M, Murakami K, Ishihama A, Adhya S (1997) Repression and activation of promoter-bound RNA polymerase activity by Gal repressor. *J Mol Biol* 272:293–300
- Craigie R, Mizuuchi K (1985) Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. *Cell* 41:867–876
- Dandaneil G, Hammer K (1985) Two operator sites separated by 599 base pairs are required for *deoR* repression of the *deo* operon of *Escherichia coli*. *EMBO J* 4:3333–3338
- Dickerson RE (1998) DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res* 26:1906–1926

- Drlica K (1987) The nucleoid. In: Neidhardt FC, Ingraham JL, Low KB, Magasanik B, Schaechter M, Umberger HE (eds) *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington, DC, pp 91–103
- Drlica K, Rouvière-Yaniv J (1987) Histone-like proteins of bacteria. *Microbiol Rev* 51:301–319
- Dunn TM, Schleif R (1984) Deletion analysis of the *Escherichia coli* P_C and P_{BAD} promoters. *J Mol Biol* 180:201–204
- Echols H, Reznicek J, Adhya S (1963) Complementation, recombination, and suppression in galactose negative mutants of *E. coli*. *Proc Natl Acad Sci USA* 50:286–293
- Friedman AM, Fischmann TO, Steitz TA (1995) Crystal structure of *lac* repressor core tetramer and its implications for DNA looping. *Science* 268:1721–1727
- Geanacopoulos M, Vasmatzis G, Lewis DE, Roy S, Lee B, Adhya S (1999) GalR mutants defective in repressosome formation. *Genes Dev* 13:1251–1262
- Geanacopoulos M, Vasmatzis G, Zhurkin VB, Adhya S (2001) Gal repressosome contains an antiparallel DNA loop. *Nat Struct Biol* 8:432–436
- Griffith J, Hochschild A, Ptashne M (1986) DNA loops induced by cooperative binding of lambda repressor. *Nature* 322:750–752
- Grove A, Galeone A, Mayol L, Geiduschek EP (1996) Localized DNA flexibility contributes to target site selection by DNA-bending proteins. *J Mol Biol* 260:120–125
- Haber R, Adhya S (1988) Interaction of spatially separated protein-DNA complexes for control of gene expression: operator conversions. *Proc Natl Acad Sci USA* 85:9683–9687
- Hahn S, Hendrickson W, Schleif R (1986) Transcription of *Escherichia coli* *ara* *in vitro*. The cyclic AMP receptor protein requirement for P_{BAD} induction that depends on the presence and orientation of the *araO2* site. *J Mol Biol* 188:355–367
- Hochschild A, Ptashne M (1986) Homologous interactions of lambda repressor and lambda Cro with the lambda operator. *Cell* 44:925–933
- Irani MH, Orosz L, Adhya S (1983) A control element within a structural gene: the *gal* operon of *Escherichia coli*. *Cell* 32:783–788
- Kabsch W, Sander C, Trifonov EN (1982) The ten helical twist angles of B-DNA. *Nucleic Acids Res* 10:1097–1104
- Kano Y, Wada M, Nagase T, Imamoto F (1986) Genetic characterization of the gene *hupB* encoding the HU-1 protein of *Escherichia coli*. *Gene* 45:37–44
- Kano Y, Osato K, Wada M, Imamoto F (1987) Cloning and sequencing of the HU-2 gene of *Escherichia coli*. *Mol Gen Genet* 209:408–410
- Kar S, Adhya S (2001) Recruitment of HU by piggyback: a special role of GalR in repressosome assembly. *Genes Dev* 15:2273–2281
- Kobryn K, Lavoie BD, Chaconas G (1999) Supercoiling-dependent site-specific binding of HU to naked Mu DNA. *J Mol Biol* 289:777–784
- Kramer H, Niemoller M, Amouyal M, Revet B, von Wilcken-Bergmann B, Muller-Hill B (1987) *lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators. *EMBO J* 6:1481–1491
- Kramer H, Amouyal M, Nordheim A, Muller-Hill B (1988) DNA supercoiling changes the spacing requirement of two *lac* operators for DNA loop formation with *lac* repressor. *EMBO J* 7:547–556
- Kuhnke G, Krause A, Heibach C, Gieske U, Fritz HJ, Ehring R (1986) The upstream operator of the *Escherichia coli* galactose operon is sufficient for repression of transcription initiated at the cyclic AMP-stimulated promoter. *EMBO J* 5:167–173
- Lavoie BD, Chaconas G (1993) Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. *Genes Dev* 7:2510–2519
- Lavoie BD, Shaw GS, Millner A, Chaconas G (1996) Anatomy of a flexer-DNA complex inside a higher-order transposition intermediate. *Cell* 85:761–771
- Law SM, Bellomy GR, Schlax PJ, Record MT Jr (1993) *In vivo* thermodynamic analysis of repression with and without looping in *lac* constructs. Estimates of free and local *lac* repressor concentrations and of physical properties of a region of supercoiled plasmid DNA *in vivo*. *J Mol Biol* 230:161–173

- Lee DH, Schleif RF (1989) In vivo DNA loops in *araCBAD*: size limits and helical repeat. *Proc Natl Acad Sci USA* 86:476–480
- Lewis DE (2003) Identification of promoters of *Escherichia coli* and phage in transcription section plasmid pSA850. *Methods Enzymol* 370:618–645
- Lewis DE, Adhya S (2002) In vitro repression of the *gal* promoters by GalR and HU depends on the proper helical phasing of the two operators. *J Biol Chem* 277:2498–2504
- Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, Schumacher MA, Brennan RG, Lu P (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271:1247–1254
- Lewis DE, Geanakopulos M, Adhya S (1999) Role of HU and DNA supercoiling in transcription repression: specialized nucleoprotein repression complex at *gal* promoters in *Escherichia coli*. *Mol Microbiol* 31:451–461
- Lia G, Bensimon D, Croquette V, Allemand JF, Dunlap D, Lewis DE, Adhya S, Finzi L (2003) Supercoiling and denaturation in Gal repressor/heat unstable nucleoid protein (HU)-mediated DNA looping. *Proc Natl Acad Sci USA* 100:11373–11377
- Lia G, Semsey S, Lewis DE, Adhya S, Bensimon D, Dunlap D, Finzi L (2008) The antiparallel loops in *gal* DNA. *Nucleic Acids Res* 36:4204–4210
- Lobell RB, Schleif RF (1990) DNA looping and unlooping by AraC protein. *Science* 250:528–532
- Majumdar A, Adhya S (1984) Demonstration of two operator elements in *gal*: in vitro repressor binding studies. *Proc Natl Acad Sci USA* 81:6100–6104
- Majumdar A, Adhya S (1987) Probing the structure of *gal* operator-repressor complexes. Conformation change in DNA. *J Biol Chem* 262:13258–13262
- Majumdar A, Rudikoff S, Adhya S (1987) Purification and properties of Gal repressor:pL-*galR* fusion in pKC31 plasmid vector. *J Biol Chem* 262:2326–2331
- Mandal N, Su W, Haber R, Adhya S, Echols H (1990) DNA looping in cellular repression of transcription of the galactose operon. *Genes Dev* 4:410–418
- Mossing MC, Record MT Jr (1986) Upstream operators enhance repression of the *lac* promoter. *Science* 233:889–892
- Muller J, Oehler S, Muller-Hill B (1996) Repression of *lac* promoter as a function of distance, phase and quality of an auxiliary *lac* operator. *J Mol Biol* 257:21–29
- Muller J, Barker A, Oehler S, Muller-Hill B (1998) Dimeric *lac* repressors exhibit phase-dependent co-operativity. *J Mol Biol* 284:851–857
- Murphy LD, Zimmerman SB (2000) Multiple restraints to the unfolding of spermidine nucleoids from *Escherichia coli*. *J Struct Biol* 132:46–62
- Musso RE, Di Lauro R, Adhya S, de Crombrughe B (1977) Dual control for transcription of the galactose operon by cyclic AMP and its receptor protein at two interspersed promoters. *Cell* 12:847–854
- Oberto J, Drlica K, Rouvière-Yaniv J (1994) Histones, HMG, HU, IHF: meme combat. *Biochimie* 76:901–908
- Oehler S, Eismann ER, Kramer H, Muller-Hill B (1990) The three operators of the *lac* operon cooperate in repression. *EMBO J* 9:973–979
- Pennington MR (2006) Sigma coupling to photons: hidden scalar in $\gamma\gamma \rightarrow \pi^0\pi^0$. *Phys Rev Lett* 97:011601
- Perez N, Rehault M, Amouyal M (2000) A functional assay in *Escherichia coli* to detect non-assisted interaction between galactose repressor dimers. *Nucleic Acids Res* 28:3600–3604
- Pettijohn DE (1988) Histone-like proteins and bacterial chromosome structure. *J Biol Chem* 263:12793–12796
- Pettijohn DE (1996) The nucleoid. In: Curtis R, Ingraham JL, Lin ECC, Low KB, Magasanik B, and Reznikoff WS et al. (eds) *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, 2nd edn. American Society for Microbiology, Washington, DC, pp 158–166
- Pontiggia A, Negri A, Beltrame M, Bianchi ME (1993) Protein HU binds specifically to kinked DNA. *Mol Microbiol* 7:343–350

- Ptashne M (1986) Gene regulation by proteins acting nearby and at a distance. *Nature* 322:697–701
- Reitzer LJ, Magasanik B (1986) Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. *Cell* 45:785–792
- Rhodes D, Klug A (1980) Helical periodicity of DNA determined by enzyme digestion. *Nature* 286:573–578
- Rice PA, Yang S, Mizuuchi K, Nash HA (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* 87:1295–1306
- Rouvière-Yaniv J (1978) Localization of the HU protein on the *Escherichia coli* nucleoid. *Cold Spring Harb Symp Quant Biol* 42(Pt 1):439–447
- Rouvière-Yaniv J, Gros F (1975) Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*. *Proc Natl Acad Sci USA* 72:3428–3432
- Roy S, Semsey S, Liu M, Gussin GN, Adhya S (2004) GalR represses *galP1* by inhibiting the rate-determining open complex formation through RNA polymerase contact: a GalR negative control mutant. *J Mol Biol* 344:609–618
- Roy S, Dimitriadis EK, Kar S, Geanakopoulos M, Lewis MS, Adhya S (2005) Gal repressor-operator-HU ternary complex: pathway of repressosome formation. *Biochemistry* 44:5373–5380
- Schleif R (1987) Gene regulation: why should DNA loop? *Nature* 327:369–370
- Schumacher MA, Choi KY, Zalkin H, Brennan RG (1994) Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science* 266:763–770
- Semsey S, Geanakopoulos M, Lewis DE, Adhya S (2002) Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *EMBO J* 21:4349–4356
- Semsey S, Tolstorukov MY, Virnik K, Zhurkin VB, Adhya S (2004) DNA trajectory in the Gal repressosome. *Genes Dev* 18:1898–1907
- Shapiro JA, Adhya SL (1969) The galactose operon of *E. coli* K-12. II. A deletion analysis of operon structure and polarity. *Genetics* 62:249–264
- Shore D, Baldwin RL (1983) Energetics of DNA twisting I. Relation between twist and cyclization probability. *J Mol Biol* 170:957–981
- Surette MG, Buch SJ, Chaconas G (1987) Transpososomes: stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA. *Cell* 49:253–262
- Tanaka H, Goshima N, Kohno K, Kano Y, Imamoto F (1993) Properties of DNA-binding of HU heterotypic and homotypic dimers from *Escherichia coli*. *J Biochem* 113:568–572
- Tullius TD, Dombroski BA (1985) Iron(II) EDTA used to measure the helical twist along any DNA molecule. *Science* 230:679–681
- Virnik K, Lyubchenko YL, Karymov MA, Dahlgren P, Tolstorukov MY, Semsey S, Zhurkin VB, Adhya S (2003) “Antiparallel” DNA loop in *gal* repressosome visualized by atomic force microscopy. *J Mol Biol* 334:53–63
- von Wilcken-Bergmann B, Muller-Hill B (1982) Sequence of *galR* gene indicates a common evolutionary origin of *lac* and *gal* repressor in *Escherichia coli*. *Proc Natl Acad Sci USA* 79:2427–2431
- Wang JC (1979) Helical repeat of DNA in solution. *Proc Natl Acad Sci USA* 76:200–203
- Wang JC, Giaever GN (1988) Action at a distance along a DNA. *Science* 240:300–304
- Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM (1998) Force and velocity measured for single molecules of RNA polymerase. *Science* 282:902–907
- Weickert MJ, Adhya S (1992) A family of bacterial regulators homologous to Gal and Lac repressors. *J Biol Chem* 267:15869–15874
- Yang B, Larson TJ (1996) Action at a distance for negative control of transcription of the *glpD* gene encoding sn-glycerol 3-phosphate dehydrogenase of *Escherichia coli* K-12. *J Bacteriol* 178:7090–7098

Chapter 18

Transcriptional Regulation

by Nucleoid-Associated Proteins at Complex Promoters in *Escherichia coli*

Douglas F. Browning, David C. Grainger, Meng Xu,
and Stephen J.W. Busby

Abstract The expression of different *Escherichia coli* transcription units is tightly regulated at the level of transcription initiation. Promoter strength is fixed by DNA sequence elements, and changes in promoter activity are primarily modulated by a combination of sigma factors and transcription factors, whose activities are controlled by the growth environment. These factors all operate in the context of bacterial chromatin which plays a key role in the expression of many transcription units. Here we describe how IHF and FIS intervene directly at some complex *Escherichia coli* promoters to bring about different regulatory outcomes. At the *nir* operon promoter, the binding of IHF and FIS together makes expression co-dependent on two transcription activators that are triggered by two different environmental signals. We discuss three different mechanisms by which FIS represses promoter activity, thereby down-regulating gene expression during rapid growth. At the *nrf* operon promoter, FIS behaves as a conventional repressor, at the *ogt* and *acs* promoters, FIS displaces the essential activator, whilst, at the *dps* promoter, FIS jams RNA polymerase containing σ^{70} in an inactive complex. In each of the three cases, derepression occurs when FIS levels drop, as cell growth slows in response to nutrient limitation. Genomic studies of the distribution of IHF and FIS across the *Escherichia coli* chromosome suggest that they intervene at many intergenic regulatory regions, and that there may be little or no distinction between some nucleoid-associated proteins and transcription factors.

Keywords Nucleoid • nucleoid-associated proteins • transcription initiation • promoter • RNA polymerase • activation • repression • integration host factor (IHF) • factor for inversion stimulation (FIS) • histone-like nucleoid structuring protein (H-NS)

D.F. Browning, D.C. Grainger, M. Xu, and S.J.W. Busby (✉)
School of Biosciences, University of Birmingham, Birmingham, B15 2TT, UK
e-mail: s.j.w.busby@bham.ac.uk

18.1 Transcription Regulation in *Escherichia coli*

Gene expression in *E. coli* is primarily controlled at the level of transcription initiation, the point at which RNA synthesis begins. The enzyme responsible for RNA synthesis is RNA polymerase and, predictably, it is the target for many regulatory factors. It is estimated that there are ~3,000–8,000 RNA polymerase molecules per *E. coli* K-12 cell, according to growth conditions, which must be distributed between ~2,500 transcription units (Ishihama 1997; Karp et al. 2007). Everything we know about bacterial transcription tells us that this distribution is not even and, thus, the cell has to regulate the binding of RNA polymerase across its chromosome prudently.

The potential activity of any promoter is set by its DNA sequence elements whilst changes in response to the environment are principally mediated by transcription factors and sigma factors (Miroslavova and Busby 2006; Browning and Busby 2004). The *E. coli* genome encodes over 250 transcription factors that exert their effects by binding at specific promoters and activating or repressing transcription (Perez-Rueda and Collado-Vides 2000). The activities of most transcription factors are regulated in response to environmental cues, usually by ligand binding, by covalent modification, or by changes in their level. A small number of transcription factors, termed “global” regulators (e.g. the cyclic AMP receptor protein, CRP), influence the expression of a large number of transcription units. Conversely, a large number of “specific” transcription factors (e.g. the Lac repressor) each affect the expression of a small number of transcription units. The expression of many transcription units is regulated by a combination of both global and specific transcription factors and this allows bacteria to differentially regulate the gene expression in response to combinations of different environmental stimuli.

The *E. coli* genome encodes seven sigma factors that exert their effects by binding to RNA polymerase, thereby steering it to specific subsets of promoters. Sigma factors thus play a pivotal role in managing the chromosome-wide distribution of the transcriptional machinery (Helmann and Chamberlin 1988; Murakami and Darst 2003). *E. coli*, like all bacteria, contains one major σ factor (σ^{70}), responsible for the recognition of most promoters. Each of the six alternative σ factors is responsible for transcription of a subset of genes, usually in response to a stress. Thus, for example, the stationary phase σ factor (σ^{38}) controls the expression of many proteins needed for the long-term survival of non-growing cells (Ishihama 1997).

Some promoters are active in the absence of additional factors and when the genes under their control are not required, they are silenced by transcription repressors. Repressor proteins reduce transcription initiation at target promoters and the textbook view is that this is simple to understand. Thus, at some promoters, a single repressor is involved and its binding prevents promoter recognition by RNA polymerase (Choy and Adhya 1996). In these instances, the repressor binding site is located at, or close to, the core promoter elements. Note that, in some cases, the repressor may not prevent binding of RNA polymerase but, rather, interferes with post-recruitment steps in transcription initiation, sometimes ‘jamming’ RNA polymerase and preventing it from initiating or elongating a transcript (Rojo 2001). At some other promoters, multiple repressor molecules bind to promoter-distal sites, and repression is caused by DNA looping, which shuts off transcription initiation within the looped domain (Semsey et al. 2005).

Most promoters lack a good match to the consensus elements for RNA polymerase binding, and many of these require ancillary proteins, known as transcription activators, to function. At some promoters, activation of transcription is simple, and involves the action of a single activator (Busby and Ebright 1994; Rhodius and Busby 1998). Two general mechanisms are used for ‘simple’ activation. In most cases, the activator binds to a target located immediately upstream of the promoter elements and recruits RNA polymerase by directly interacting with a target usually in the RNA polymerase α or σ subunits. In a small number of cases, the activator alters the conformation of the target promoter, to enable the interaction of RNA polymerase with the promoter -10 and -35 elements (Brown et al. 2003).

Most bacterial promoters are regulated by more than one transcription factor and this permits regulatory input from multiple environmental cues (Martinez-Antonio and Collado-Vides 2003). Different mechanisms have evolved for integrating the effects of different transcription factors at such complex promoters. Thus, at promoters that are co-dependent on two or more activators, several complicated mechanisms have been discovered, involving the repositioning of one activator by another, independent activator-RNA polymerase contacts, or anti-repression by an activator (Barnard et al. 2004; Browning and Busby 2004). For mechanisms involving repositioning, the role of the secondary activator is to reposition the primary activator from a location where it is unable to activate transcription, to a location where it can activate transcription. This repositioning can involve either shifting the primary activator from one DNA site to another, or altering the conformation of the DNA to allow the primary activator to interact with RNA polymerase. For example, at the *E. coli malK* promoter, CRP repositions MalT from a non-productive high affinity binding site to a location where it can interact with RNA polymerase (Richet et al. 1991). In contrast, at the *narG* promoter, IHF induces a DNA bend that permits upstream-bound NarL to activate transcription (Schroder et al. 1993). A different mechanism operates at promoters where two activators must each make contact with RNA polymerase for transcription. In these cases, one activator-RNA polymerase contact is insufficient. In most cases studied to date, the two activators bind independently at the target promoter. For example, at the *E. coli ansB* promoter, activation depends on CRP and FNR, bound independently at positions -91 and -41 respectively, making separate contacts with RNA polymerase (Scott et al. 1995). Finally, in some cases, the role of the second activator is not to activate directly, but rather to prevent the action of a repressor that is interfering with the function of the primary activator. In these cases, the second activator behaves as an anti-repressor rather than a true activator (Browning and Busby 2004).

18.2 Nucleoid-Associated Proteins Can Participate at Promoters

The mechanisms of transcriptional regulation, described above, all operate in the context of the bacterial chromosome, which is folded into a compact structure, known as the nucleoid. Clearly, the distribution of RNA polymerase between different regulatory regions will be affected by this compaction. Most attention has

focused on the nucleoid-associated proteins that, alongside supercoiling and macromolecular crowding, induced by RNA and other proteins, are involved in maintaining compaction (Thanbichler et al. 2005; Dame 2005). In a seminal study, Akira Ishihama and colleagues defined 12 different nucleoid-associated proteins: HU, H-NS, StpA, FIS, IHF, CbpA, CbpB, Dps, Lrp, DnaA, Hfq and IciA (Talukder et al. 1999). Most of these induce conformational changes in DNA upon binding in vitro (e.g. bending or bridging) but, in many cases, the functional significance of the changes in vivo are unclear. All of these proteins bind DNA non-specifically, but some also have higher affinity for specific sequences. In their study, Ishihama and colleagues raised antibodies against each of the 12 nucleoid-associated proteins, and their levels were quantified during different stages of cell growth. Whilst the levels of each different nucleoid-associated protein varied throughout growth, each of the nucleoid-associated proteins was present at over 10,000 copies per cell at some stage (with the exception of Lrp, DnaA and IciA). The most dramatic changes during cell growth were found with FIS, Dps and CbpA. Rapidly growing cells contain over 50,000 molecules of FIS and near zero levels of Dps and CbpA. When cell growth slows in post-exponential phase, FIS levels drop and Dps levels rise. When cells stop growing and arrive in stationary phase, FIS levels are near zero and maximum amounts of Dps (>20,000 molecules) are found. At this point, CbpA is induced and reaches ~12,000 molecules per cell. Figure 18.1 illustrates the changes in the levels of several of the nucleoid-associated proteins as *E. coli* cells pass from exponential to stationary phase and summarises some of the principal properties of each factor.

An important point is that many of the abundant nucleoid-associated proteins also act as transcription factors (McLeod and Johnson 2001), and many of these functions are described in this section of this book. These proteins regulate transcription initiation by binding to specific sites at target promoters. Thus both IHF and FIS can function to activate transcription initiation, behaving as ‘simple’ transcription factors by recruiting RNA polymerase to target promoters. However, both IHF and FIS have also been reported to activate transcription by binding upstream of target promoters and inducing conformation changes in the downstream sequence. Thus, IHF binding upstream of the *ilvG* promoter alters the conformation of the *ilvG* -10 element, thereby activating transcription (Sheridan et al. 2001), and similar effects have been reported during the activation of the *leuV* promoter by FIS (Opel et al. 2004). Opening of the DNA duplex by negative supercoiling plays a key role in this process and is often referred to as supercoiling-induced DNA duplex destabilization.

At other promoters, nucleoid-associated proteins facilitate the action of activators or repressors. Hence at many promoters whose activity depends on a direct interaction between an upstream-bound transcription activator and RNA polymerase containing σ^{54} , DNA bending induced by IHF binding between the activator and RNA polymerase is essential to ensure that the interaction occurs (Wigneshweraraj et al. 2008). Similarly, HU binding is required for GalR-dependent repression of *galETK* expression. In this case, HU-induced DNA bending permits an interaction between upstream-bound and downstream-bound GalR dimers that is essential for efficient repression (Roy et al. 2005) (Chapter 17).

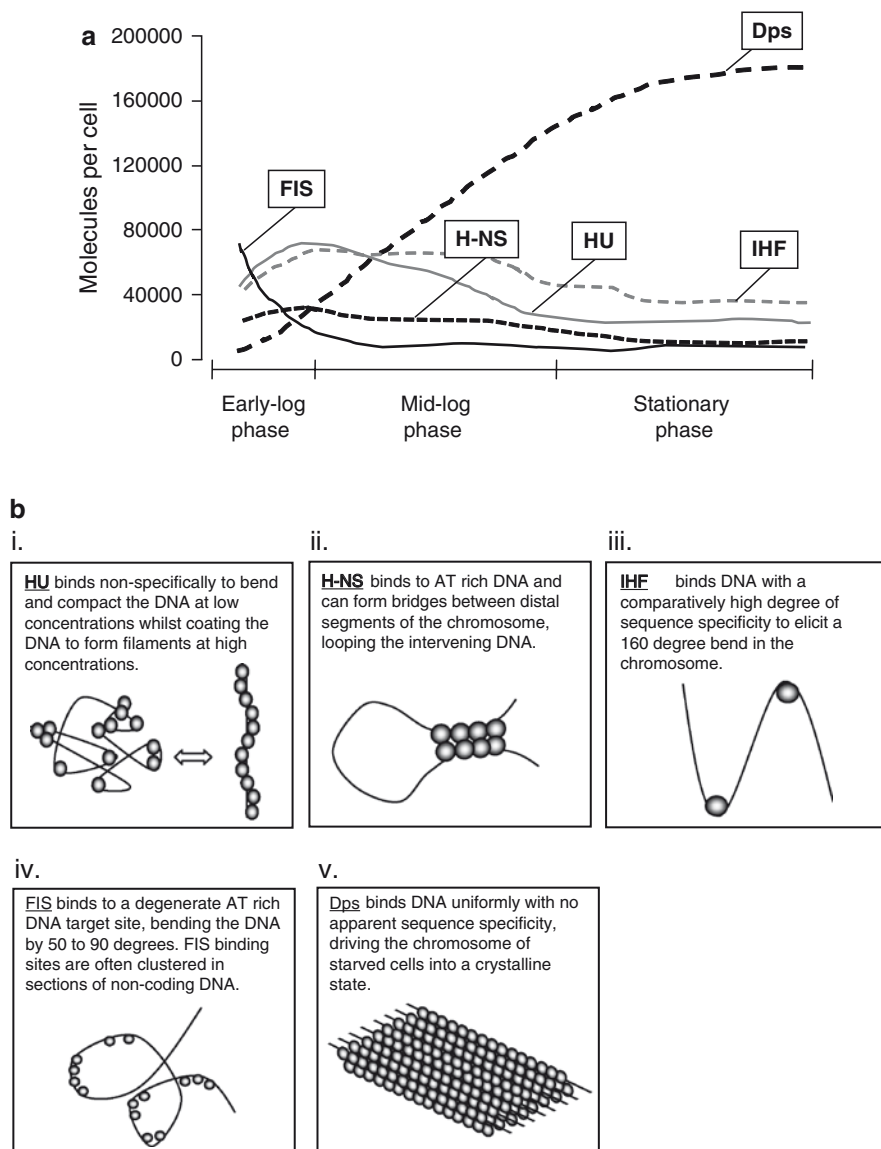


Fig. 18.1 The principal nucleoid-associated proteins of *E. coli*. (a) The figure depicts the fluctuations in the levels of the major nucleoid-associated proteins that occur with changes in growth phase (data taken from Talukder et al. 1999). (b) Panels i, ii, iii, iv and v illustrate the main impact of HU, H-NS, IHF, FIS and Dps proteins respectively on chromosome structure. The various nucleoid-associated proteins are shown as spheres and DNA is shown as a solid line

18.3 Repression and Antirepression at the *Escherichia coli nir* Operon Promoter

Nucleoid-associated proteins play quite subtle roles at some complex promoters and perhaps the best example of this is found at the *E. coli nir* operon regulatory region. The *nir* operon encodes a cytoplasmic nitrite reductase that catalyses the NADH-dependent reduction of nitrite ions to ammonia (Harborne et al. 1992). Transcription from a single startpoint is controlled by a promoter upstream of the *nirB* gene (Jayaraman et al. 1988). At this promoter, H-NS is an overall repressor, whereas FIS and IHF function in concert to confer codependence on two activators (Browning et al. 2000).

Early studies had shown that the *nir* promoter is optimally active when cells are grown in anaerobic conditions in the presence of nitrite or nitrate ions, and also that higher activities are found in rich media (Bell et al. 1990; Page et al. 1990). Induction in anaerobic conditions is due to the activity of FNR, a global transcription activator responsible for the induction of over 100 different transcription units in *E. coli* in response to low oxygen levels (Browning et al. 2002a). FNR dimerisation, and hence specific binding at target promoters, requires the formation of an iron-sulphur cluster that is destroyed by oxygen. However as oxygen levels decrease, the iron-sulphur cluster forms, FNR dimerises and binds to target sites, and is then able to activate transcription. In most cases, including the *nir* promoter, the FNR dimer binds to a target near position -41 and functions as a simple activator of transcription initiation (Wing et al. 1995; Fig. 18.2).

In addition to being dependent on FNR, expression from the *nir* promoter is dependent on activation by NarL (or its homologue, NarP), that are response regulators, which are both activated by nitrate or nitrite ions via the membrane bound NarX and NarQ sensor kinases. It is this dependence that ensures that *nir* operon induction is coupled to the presence of nitrate or nitrite ions in the environment, as well as the lack of oxygen (Tyson et al. 1993, 1994). NarL binds as a dimer to a DNA site just upstream of the DNA site for FNR at the *nir* promoter (Fig. 18.2) and has no effect on FNR binding. Since FNR is perfectly able to function as an activator alone, this raises the question of why NarL is needed. The explanation for this follows from the observation that the sequences upstream of the DNA site for FNR carry targets for FIS at position -142 and for IHF at positions -88 and -115 (Browning et al. 2000, 2004a). Genetic experiments show that FNR-dependent activation is suppressed by the binding together of FIS at position -142 and IHF at position -88. In vitro experiments with purified components demonstrate that FNR is able to activate open complex formation at the *nir* promoter and that the complex is destabilized by binding of FIS and IHF. The suppression mediated by FIS and IHF is relieved upon binding of NarL, which displaces IHF from position -88 (Fig. 18.2). Thus here, the role of the second activator, NarL, is not to activate directly, but rather to prevent FIS and IHF from interfering with the function of the primary activator (FNR). Hence, point mutations that destroy the DNA sites for FIS and IHF at positions -142 and -88 respectively, release the requirement for NarL, and convert the *nir* promoter into a simple FNR-dependent promoter (Wu et al. 1998).

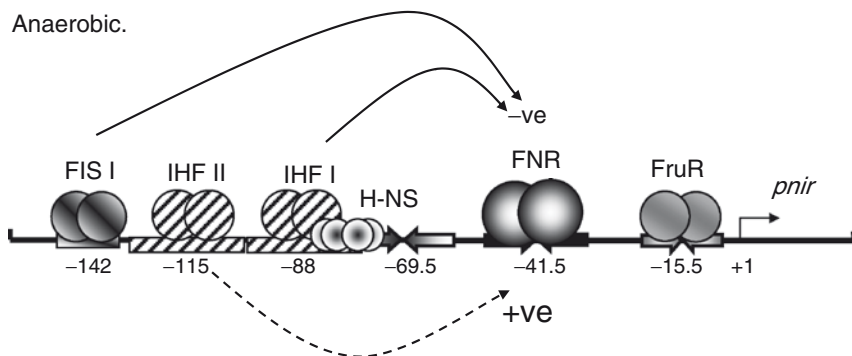
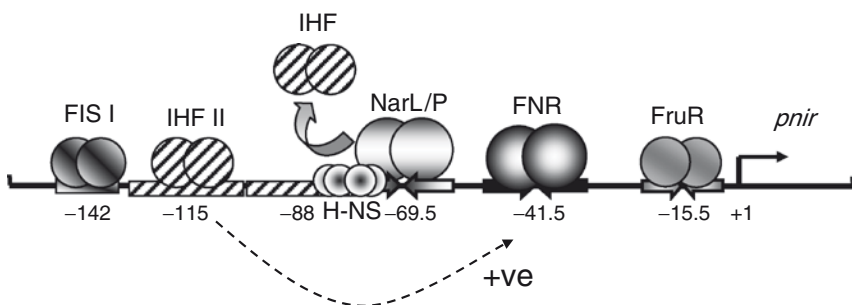
a Anaerobic.**b Anaerobic plus NO_2^- / NO_3^- .**

Fig. 18.2 Transcription regulation at the *E. coli nir* operon promoter. **(a)** Anaerobic conditions. Transcription initiation is dependent on FNR. The binding of FIS to FIS I and IHF to IHF I inhibits FNR-activated transcription ($-ve$), whilst IHF binding to the lower affinity IHF II site stimulates transcription ($+ve$). The mechanisms by which upstream bound FIS and IHF modulate FNR-dependent transcription initiation at the *nir* promoter are not understood. In addition, the binding of H-NS and FruR can down-regulate transcription initiation at the *nir* promoter. **(b)** Anaerobic conditions plus nitrite or nitrate. The binding of NarL (or its homologue, NarP) displaces IHF from IHF I, thus counteracting the repression mediated by IHF and FIS, and enabling maximal FNR-dependent transcription. In contrast, NarL and NarP have little effect on repression by H-NS and FruR, whose effects are modulated by temperature and nutrient richness respectively

Further analysis has revealed a complication: IHF binding to a weaker second site at position -115 promotes rather than suppresses FNR-dependent transcription activation (Browning et al. 2004a). Thus, IHF bound at two adjacent sites (at positions -88 and -115) have opposite effects on *nir* promoter activity (Fig. 18.2). Interestingly, the relative IHF binding affinities at the two sites differ in diverse enteric bacteria and this sets the basal level of FNR-dependent activation in the absence of NarL (Browning et al. 2008). Hence the basal NarL-independent activity of the *nir* promoter is increased by mutations that improve IHF binding at position -115 and decreased by mutations that destroy binding (Browning et al. 2004a). A second complication is that H-NS globally down-regulates *nir* promoter activity and, thus, three different nucleoid-associated proteins participate directly in

regulating the *E. coli nir* operon promoter (Browning et al. 2000). The effects of H-NS are due to a specific single target (Fig. 18.2). Although this target overlaps the DNA site for NarL/NarP, H-NS binding does not appear to interfere with the actions of NarL and NarP.

Although activity of the *nir* promoter is primarily controlled by lack of oxygen and the presence of nitrite or nitrate ions, it is also sensitive to the richness of the growth medium, which provides a third environmental cue (Page et al. 1990). Hence, in nutrient-poor media, *nir* promoter activity is repressed irrespective of anaerobiosis or nitrite and nitrate ions. Nucleoid-associated proteins play little or no role in this regulation but, rather, it is due to the binding of FruR to a target site centered at position -15.5 (Tyson et al. 1997). This target site is just upstream of the *nir* promoter -10 element and hence FruR binding overrides all other regulation simply by blocking RNA polymerase binding (Fig. 18.2). FruR (also known as Cra) is a member of the *lac* repressor family of transcription factors and is a global regulator whose binding at target sites is modulated by fructose diphosphate (Saier and Ramseier 1996; Shimada et al. 2005). Thus FruR represses *nir* expression when nutrients are in short supply and glycolytic flux is low but repression is relieved in nutrient-rich growth media.

18.4 The *nrf* Operon Promoter: Variation on a Theme

At the *E. coli nir* operon promoter, the main role of FIS and IHF is to confer co-dependence of the promoter on two activators, FNR and NarL (or NarP), thereby coupling *nir* operon expression to two physiological signals. FIS appears to play only a supporting role to IHF, which is the primary repressor of FNR-dependent activation, and, hence, the primary target for NarL/NarP (Browning et al. 2000, 2004a). FruR is responsible for the changes in *nir* promoter activity that occur in response to nutrient richness. Thus, fluctuations in FIS levels, in response to changes in growth rate and nutrient abundance, have little or no consequence. An interesting variation of this theme is found at the regulatory region of the *E. coli nrf* operon, which encodes a periplasmic formate-dependent nitrite reductase (Hussain et al. 1994; Tyson et al. 1994). Transcription from a single startpoint is controlled by a promoter upstream of the *nrfA* gene and codependence on two physiological signals, anaerobiosis and nitrate/nitrite ions, is conferred by IHF alone, without FIS (Browning et al. 2002b). The role of FIS is to repress the promoter in response to nutrient richness (Browning et al. 2005). Thus, whilst the *nir* operon promoter is induced in rich media, the *nrf* operon promoter is repressed and this repression is due to FIS binding to a target that overlaps the promoter -10 element (Fig. 18.3). Hence FruR and FIS bind to similar locations at the *nir* and *nrf* promoters respectively, thereby shutting down the promoters, respectively in response to nutrient scarcity and abundance, and overriding the effects of other transcription factors bound further upstream (compare Figs. 18.2 and 18.3).

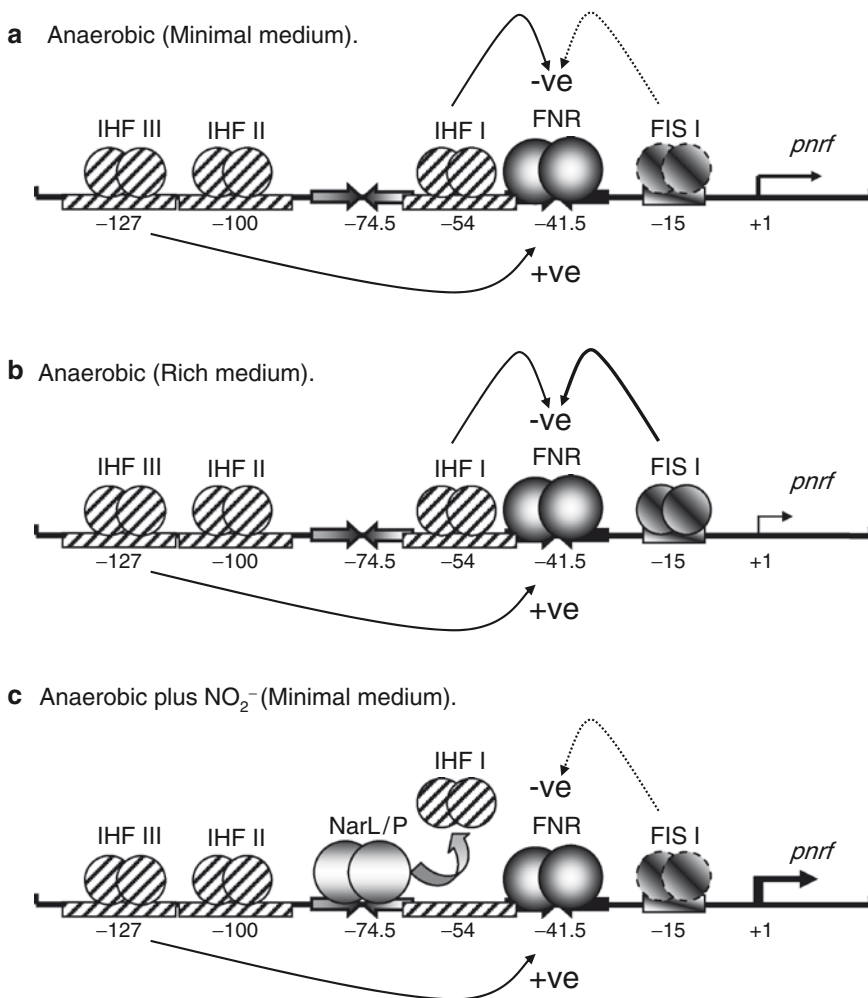


Fig. 18.3 Transcription regulation at the *E. coli nrf* operon promoter. **(a)** Minimal medium. Transcription initiation is dependent on FNR. The binding of IHF to IHF I and FIS to FIS I inhibits FNR-activated transcription (-ve), whilst occupancy of the upstream IHF III site stimulates transcription (+ve). In minimal medium FIS protein levels are low and so FIS I is not fully occupied (shown as dotted). **(b)** Rich medium. Higher levels of FIS protein lead to increased occupancy of FIS I, resulting in further repression of *nrf* operon expression. **(c)** Minimal medium with nitrite. The binding of NarL (or its homologue, NarP) displaces IHF from IHF I, counteracting IHF-mediated repression and enabling maximal FNR-dependent transcription

Early studies had shown that the *nrf* operon promoter is optimally active when cells are grown in anaerobic conditions in the presence of nitrite or nitrate ions, and is repressed in nutrient rich conditions (Page et al. 1990). As with the *nir* promoter,

the *nrf* promoter is totally dependent on FNR, which binds to a target centered near position -41 and functions as a simple activator (Tyson et al. 1994). Activation in response to nitrate and nitrite ions is dependent on NarL or NarP binding to a site at position -74.5 (in contrast to position -69.5 at the *nir* operon promoter). Studies have shown that FNR alone is sufficient to activate the *nrf* promoter fully, but that FNR-dependent activation is repressed by IHF binding to a DNA site located at position -54 (IHF site I, see Fig. 18.3). NarL or NarP binding at position -74.5 displaces IHF from site I, thus permitting FNR-dependent activation (Browning et al. 2002b, 2006). Thus the primary role of IHF binding at site I is to confer codependence of expression on two activators, and hence two physiological signals. This mechanism is very similar to the one that operates at the *nir* promoter, save that, at the *nrf* promoter, IHF binds downstream of the DNA site for NarL/NarP, and does not require support from FIS. Interestingly, as with the *nir* promoter, there are other upstream sites for IHF and the binding of IHF to one of them (IHF III) stimulates FNR-dependent activation of the *nrf* promoter (Browning et al. 2006). Hence activation of the *E. coli nir* and *nrf* promoters shares a common mechanism yet uses different arrangements of the ‘actors’ and binding sites (compare Figs. 18.2 with 18.3).

The situation at the *nrf* promoter is complicated by the divergent *acs* promoter that is located upstream and drives transcription from a startpoint located 280 base pairs upstream of the *nrf* promoter startpoint. Expression of the *acs* promoter is dependent on CRP and this activation is moderated by IHF binding at site II and site III (Beatty et al. 2003; Browning et al. 2004b; Sclavi et al. 2007). Thus, the primary roles of IHF bound at site II and site III appear to be concerned with the *acs* promoter, and the stimulation of the *nrf* promoter may be serendipitous.

As mentioned above, the most striking difference between the *E. coli nir* and *nrf* promoters is that *nrf* promoter expression is repressed in rich medium and that FIS has been relieved of its duty as a ‘helper’ for IHF and given the role of ‘master’ repressor. To do this, the DNA site for FIS at the *nrf* promoter is located at position -15 , corresponding exactly to the location of the DNA site for FruR at the *nir* promoter. Some of the supporting evidence for the effects of FIS at the *nrf* promoter is shown in Fig. 18.4. Full repression in vivo requires the *fis* gene and

Fig. 18.4 (continued) concentrations of FIS and subjected to DNase I footprinting. The concentrations of FIS used in each incubation were: lane 1, zero; lane 2, 0.06 μM ; lane 3, 0.11 μM ; lane 4, 0.22 μM ; lane 5, 0.44 μM ; lane 6, 0.88 μM . Gels were calibrated using Maxam-Gilbert ‘G + A’ sequencing reactions and relevant positions are indicated. The location of FIS I is shown by a box. (c) In vitro transcription assays using plasmid pSR/ *pnr97*, carrying the *nrf* promoter upstream of a strong transcription terminator, as template. The RNA I transcript, encoded by sequences within the pSR vector, serves as an internal control for transcript formation. The *nrf* and RNA I transcripts are indicated by arrows. The concentrations of FNR used in incubations were: lane 1, zero; lanes 2–6, 1 μM . The concentrations of FIS were: lanes 1 and 2, zero; lane 3, 0.11 μM ; lane 4, 0.22 μM ; lane 5, 0.44 μM ; lane 6, 0.88 μM

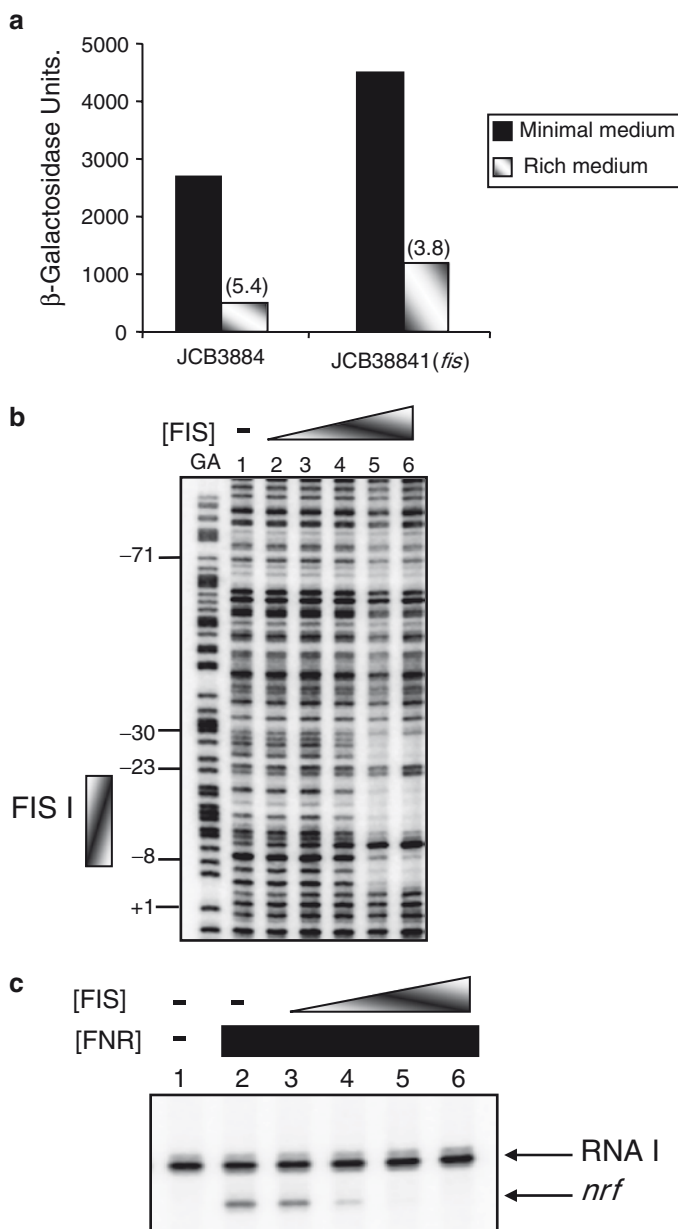


Fig. 18.4 FIS represses FNR-dependent transcription at the *E. coli nrf* promoter. **(a)** The panel shows measured β -galactosidase activities of JCB3884 (*fis*⁺) and JCB38841 (*fis*) cells carrying pRW50 (a *lac* expression vector plasmid), containing the *pnrf53/Δ87* promoter fragment (Browning et al. 2005). Cells were grown anaerobically in either minimal or rich media and β -galactosidase activities are expressed as nmol of ONPG hydrolyzed min⁻¹ mg⁻¹ dry cell mass. The fold catabolite repression for each strain is indicated in brackets. **(b)** DNase I footprint analysis of the *nrf* promoter. End-labelled *pnrf53* *Aat*II-*Hind*III fragment was incubated with increasing

an intact DNA site for FIS. Binding of FIS to this site can be demonstrated in vitro and this binding represses FNR-dependent activation of transcription in a simple assay system.

18.5 FIS Protein as a Sensor of Nutrient Abundance in *E. coli*

The intracellular level of FIS rises dramatically in response to nutrient availability and rapid growth (Ball et al. 1992). This is exploited at the *nrf* operon promoter to shut off the expression of a nitrite reductase when its role as a scavenger for nitrite in nutrient-poor conditions is not required (Wang and Gunsalus 2000). The location of the DNA site for FIS, overlapping the -10 element, suggests that FIS acts as a simple repressor, similar to many repressors, functioning by preventing RNA polymerase binding (Rojo 2001). The *E. coli* genome contains many genes whose function is coping with nutrient poor conditions and transcriptional repression in response to nutrient excess is a common feature in the regulation of the promoters of these genes. Whilst most of these effects may be due to CRP, we can expect that FIS will be involved in repression in many cases. In addition to the *nrf* promoter, three further examples, the *ogt*, *acs* and *dps* promoters have recently been reported. These cases are particularly interesting since the mechanism of repression by FIS at these promoters is more complex.

Figure 18.5 shows a sketch of the *E. coli ogt* regulatory region which contains a single promoter that is activated by NarL alone without the participation of FNR (in contrast to the *nir* and *nrf* promoters). NarL binds to two targets at position -45.5 and position -78.5 . Results illustrated in Fig. 18.6 show that *ogt* expression is induced by nitrate ions, that this induction is dependent on NarL, and that expression is repressed in rich medium. This repression is dependent on FIS that binds to a single site located at position -82 , which overlaps the upstream DNA site for NarL. The footprinting experiment illustrated in Fig. 18.6c shows that FIS binding displaces NarL, which is unsurprising since the two operator sites overlap. The simplest explanation is that repression by FIS, which is triggered by the rise in FIS levels that occur in rich medium, is due to the displacement of NarL. Thus, at the *ogt* promoter, repression by FIS is due to the displacement of an activator, rather than blocking RNA polymerase binding (as at the *nrf* promoter).

The *ogt* promoter is especially interesting as it controls the expression of an enzyme that repairs methylated DNA (Potter et al. 1987; Samson 1992). Most NarL-dependent promoters are co-dependent on FNR and encode proteins involved in nitrate and nitrite metabolism, and, to date, very few promoters have been found to be activated by NarL without FNR (Constantinidou et al. 2006; Lin et al. 2007). The fact that nitrate and nitrite metabolism generates harmful reactive nitrogen species (Taverna and Sedgwick 1996; Weiss 2006) prompts a simple explanation for our findings. Some reactive nitrogen species interact with amino acid side chains and that promotes DNA methylation. We suggest that the NarL-dependent induction

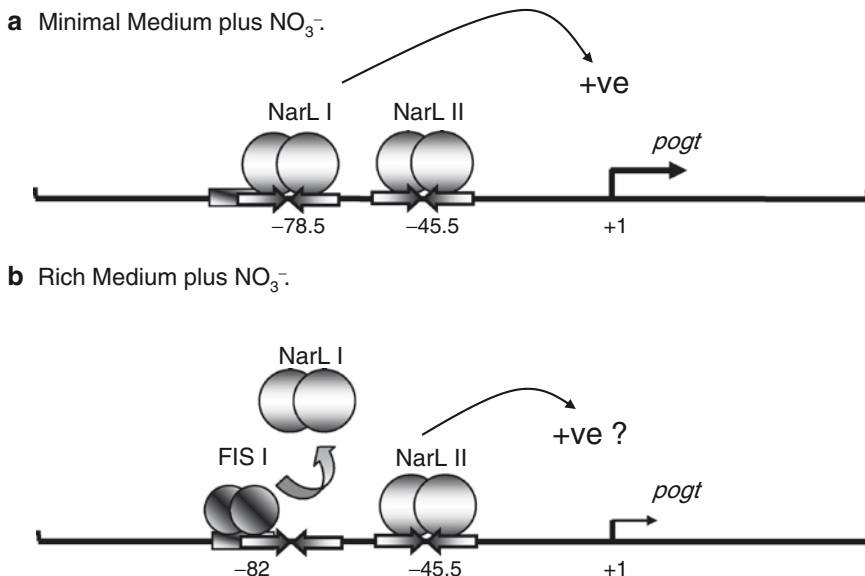


Fig. 18.5 Transcription regulation at the *E. coli ogt* promoter. (a) Minimal medium with added nitrate. The binding of NarL at the NarL I and NarL II sites activates transcription (+ve) at the *ogt* promoter. (b) Rich medium with added nitrate. FIS, binding to FIS I, displaces NarL, bound at NarL I, repressing transcription. It is not yet known if NarL bound to NarL II is able to activate transcription independently of NarL at NarL I

of *ogt* protects the cell's DNA against the mutational consequences of such reactions, and that such protection is deemed unnecessary in nutrient rich conditions, where DNA damage arising from the side products of high metabolic flux poses a far more substantial threat.

A second example of FIS-dependent repression due to displacement of an essential activator is found at the *E. coli acs* promoter, which controls expression of acetyl-coenzyme A synthetase. The *acs* promoter is divergent from the *nrf* promoter and the intergenic region contains a complicated array of binding sites for FIS and IHF that affect both promoters (Fig. 18.7). Expression from the *acs* promoter is dependent on activation by tandem-bound CRP at position -69 and -122 (Beatty et al. 2003) and is repressed by both FIS and IHF, which function by different mechanisms. FIS binding at position -59 (FIS III) and position -98 (FIS II) represses the *acs* promoter by displacing CRP from both sites. In contrast, IHF, which binds to three upstream sites, at positions -153, -180 and -226 represses the *acs* promoter without displacing CRP (Browning et al. 2004b; Sclavi et al. 2007). The combined effects of FIS and IHF ensure that *acs* expression peaks at the transition from exponential to stationary phase. Recall that acetyl-coenzyme A synthetase enables the cell to use acetate and is transiently expressed as cells enter

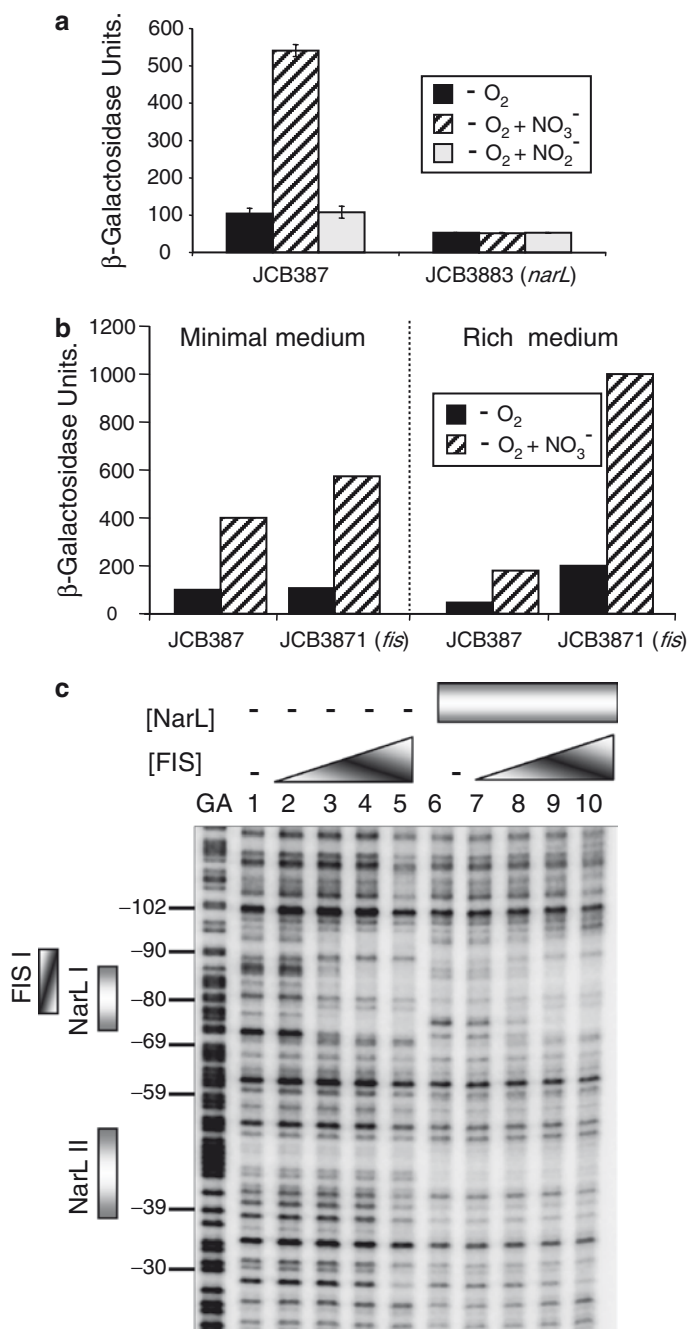


Fig. 18.6 FIS represses the *E. coli ogt* promoter in rich growth media. (a) The panel shows measured β -galactosidase activities of JCB387(*narL*⁺) and JCB3883 (*narL*) cells carrying pRW50, containing the *pgt100* promoter fragment. Cells were grown anaerobically in minimal medium and nitrate and nitrite were added to a final concentration of 20 and 2.5 mM, respectively, where indicated.

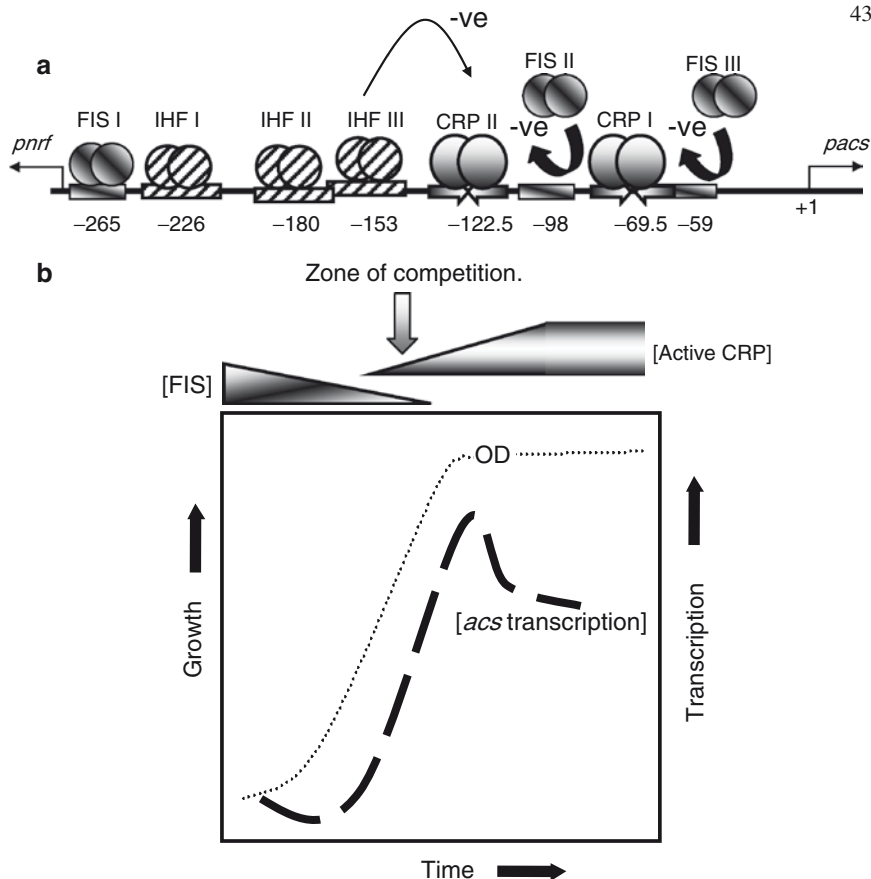


Fig. 18.7 Transcription regulation at the *E. coli acs* promoter. (a) Regulation of *pacS*. The binding of CRP to the CRP I and CRP II sites activates transcription at the *acs* promoter. In exponential phase when cellular FIS concentrations are high, FIS binding to FIS II and FIS III displaces CRP and represses transcription (-ve). The binding of IHF to three upstream sites also represses CRP-dependent transcription (-ve). (b) Modulation of CRP-dependent *acs* transcription by FIS. The figure shows how *acs* transcription alters due to bacterial growth (OD). The relative concentrations of active CRP and FIS are shown

Fig. 18.6 (continued) (b) The panel shows measured β -galactosidase activities of JCB387(*fis*⁺) and JCB3871(*fis*) cells carrying pRW50, containing the *pogt100* promoter fragment. Cells were grown anaerobically in either minimal or rich media and nitrate was added to a final concentration of 20 mM where indicated. β -galactosidase activities are expressed as nmol of ONPG hydrolyzed min⁻¹ mg⁻¹ dry cell mass and each activity is the average of three independent determinations. (c) DNase I footprint analysis of the *ogt* promoter. End-labelled *pogt100* *AatII-HindIII* fragment was incubated with increasing concentrations of FIS in combination with NarL and subjected to DNase I footprinting. The concentrations of FIS used in incubations were: lanes 1 and 6, zero; lanes 2 and 7, 0.45 μ M; lanes 3 and 8, 0.89 μ M; lanes 4 and 9, 1.8 μ M; lanes 5 and 10, 3.8 μ M. The concentrations of NarL were: lanes 1–5, zero; lanes 6–10, 3.2 μ M. Gels were calibrated using Maxam-Gilbert 'G + A' sequencing reactions and relevant positions are indicated. The locations of NarL and FIS binding sites are indicated by vertical boxes

stationary phase and take up and metabolise acetate that had been excreted during exponential growth (Wolfe 2005).

18.6 Regulation of the *Escherichia coli* *dps* Promoter

Dps is a nucleoid-associated protein that is absent in rapidly growing *E. coli* but accumulates as growth slows and cells enter stationary phase (Almiron et al. 1992). In non-growing cells, Dps becomes the most abundant nucleoid-associated protein and this is thought to be a key factor in maintaining the stationary phase folded chromosome. The expression of Dps depends on a single promoter located just upstream of the *dps* gene and accumulation of Dps requires the stationary phase σ factor, σ^{38} . The observation that the *dps* promoter can be served by RNA polymerase containing either σ^{38} or the major σ factor, σ^{70} , raises the puzzle of what prevents *dps* from being expressed in rapidly growing cells (Altuvia et al. 1994). Recent studies have shown that, just as at the *nrf* and *ogt* promoters, FIS is the key factor in repression, but that it acts by an unusual mechanism in which FIS jams RNA polymerase containing σ^{70} but not σ^{38} at the *dps* promoter (Grainger et al. 2008). Thus, in rapidly growing cells, the *dps* promoter is silenced by a ternary repression complex containing RNA polymerase with σ^{70} , FIS and promoter DNA. Remarkably, FIS has little or no effect on the activity of RNA polymerase containing σ^{38} , and hence FIS can discriminate between two different forms of RNA polymerase. This provides an efficient switch for ensuring that the *dps* promoter is silent, when FIS levels are high, but activated as FIS levels fall and σ^{38} levels rise.

As well as being repressed by FIS, the *dps* promoter is also regulated by H-NS, IHF and OxyR (Fig. 18.8). Like FIS, H-NS acts as a repressor that discriminates between RNA polymerase containing σ^{70} and σ^{38} (Grainger et al. 2008). H-NS displaces RNA polymerase containing σ^{70} from the *dps* promoter, whilst not interfering with RNA polymerase containing σ^{38} . Thus, together with FIS, H-NS confers σ factor dependence on *dps* expression. In contrast, a third nucleoid-associated protein, IHF, binds upstream of the core *dps* promoter elements and functions as an activator during σ^{38} -dependent transcription in stationary phase (Altuvia et al. 1994; Ohniwa et al. 2006). Finally a second activator, OxyR, which is triggered by oxidative stress, also binds upstream, and is responsible for transient induction of *dps* during oxidative stress in rapidly growing cells (Altuvia et al. 1994; Ohniwa et al. 2006). In these circumstances, the repression by FIS and H-NS must be overcome, but the mechanism for this is unclear at present. Panels a–c of Fig. 18.8 illustrate the molecular complexes responsible for silencing the *dps* promoter in rapidly growing cells, and activating it in response to oxidative stress in growing cells, or as cells progress to stationary phase (Schnetz 2008).

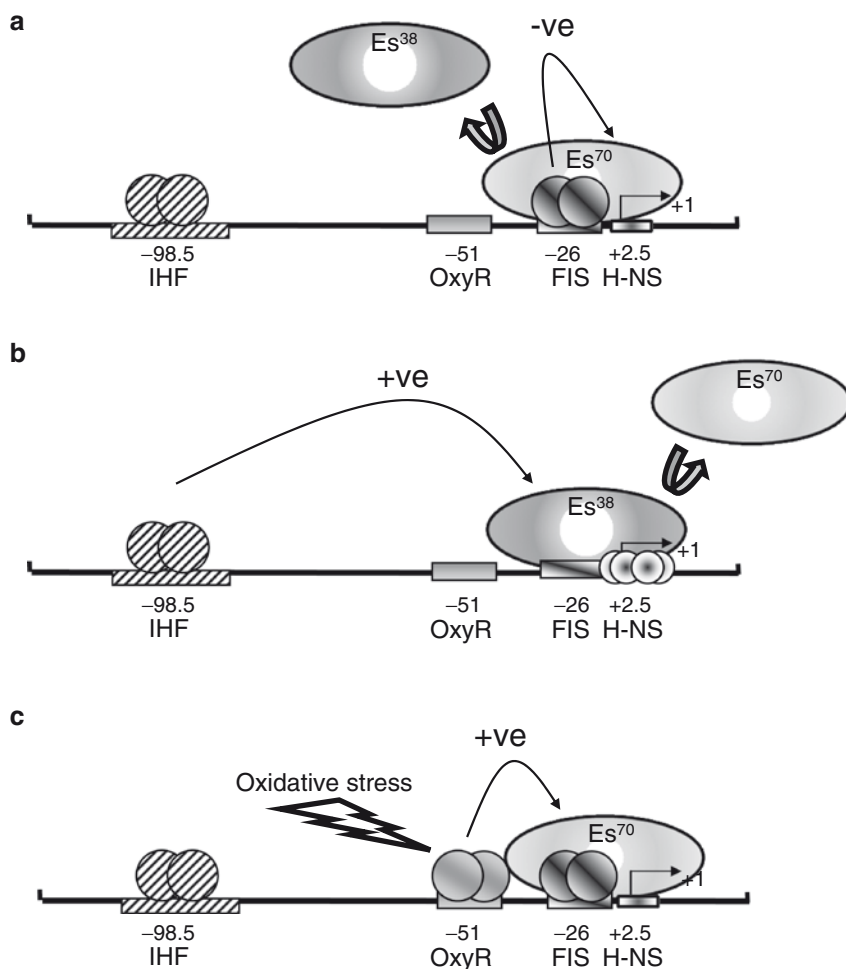


Fig. 18.8 Selective regulation of the *E. coli dps* promoter. (a) Selective repression by FIS. During rapid growth, transcription from the *dps* promoter is repressed by FIS, which binds to the promoter in unison with RNA polymerase containing σ^{70} and shuts down the promoter, blocking access by RNA polymerase containing σ^{38} . (b) Selective repression by H-NS. Binding of H-NS to the *dps* promoter blocks the binding of RNA polymerase with σ^{70} but permits binding of RNA polymerase with σ^{38} . Transcription by RNA polymerase with σ^{38} (but not with σ^{70}) can be stimulated by IHF. (c) Activation by OxyR. In response to oxidative stress, transcription from the *dps* promoter by RNA polymerase containing σ^{70} is enhanced by OxyR, which overcomes the negative effects of FIS and H-NS by a yet unknown mechanism (Schnetz 2008)

18.7 Genome-Wide Effects of FIS and IHF

The above examples underscore the versatility of FIS and IHF in moderating regulation at bacterial promoters. To gain insight into the global roles of these factors, chromatin immunoprecipitation has been exploited to find their binding locations across the whole *E. coli* K-12 chromosome (Grainger and Busby 2008). To do this, the sequence composition of DNA fragments, which had been immunoprecipitated with antisera directed against either FIS or IHF, was analysed, using high density microarrays (Grainger et al. 2006). Figure 18.9 shows a typical set of results illustrating the distribution of FIS and IHF. Each scan shows the enrichment (y-axis) for DNA sequences at particular loci (x-axis) in the immunoprecipitated DNA samples. As expected, both proteins bind at many targets. For FIS and IHF, 224 and 135 targets respectively were identified, and these include most of the previously identified targets (63 and 55 targets, respectively, listed in *EcoCyc*: Karp et al. 2007). Surprisingly, ~60% of the targets for both FIS and IHF are in intergenic regulatory regions. Since these regions cover less than 10% of the total genome, it is clear that FIS and IHF binding must be highly focused. This is unlike the situation with eucaryotic histones that bind at equal densities to both coding and non-coding targets. It is clear that, if FIS and IHF are involved in chromosome compaction, they must orchestrate this primarily by binding at intergenic regulatory regions. Analysis of the target locations revealed 54 regulatory regions where FIS and IHF both interact, including the *E. coli nir*, *nrf-acs* and *dps* operon regulatory regions.

Many authors consider the *E. coli* nucleoid-associated proteins to be different from transcription factors. However, the above analysis with FIS and IHF questions this distinction, since their binding profile resembles that of some transcription factors. Figure 18.10 shows the binding profiles of two of the best characterized *E. coli* transcription factors, CRP and FNR. As for FIS and IHF, these profiles were derived from analyzing immunoprecipitated DNA using antisera directed against purified CRP (Grainger et al. 2005) or against a FLAG tag that was attached to FNR (Grainger et al. 2007). For both transcription factors, 60–70 clear targets were identified, which correspond to both previously identified and previously unidentified targets. At most of these targets, factor binding can either up-regulate or down-regulate transcription initiation. However, at some targets, it was impossible to measure any detectable consequence on the activity of promoter activity, suggesting that there may be bona fide targets for transcription factor binding that serve no purpose, at least directly, in the modulation of gene expression (Grainger et al. 2007; Hollands et al. 2007). Note that a similar conclusion was found with the *E. coli* RutR factor (Shimada et al. 2008). Strikingly, the binding profile for CRP shows a strong background that appears to be due to its binding to many thousands of weak sites across the *E. coli* chromosome (Grainger et al. 2005). Interestingly, these sites were predicted by bioinformatic studies (Robison et al. 1998) and are consistent with the observed ‘non-specific’ binding of CRP observed in electromobility shift assays (Kolb et al. 1983). Clearly, it is unlikely that transcription is regulated from these sites, and, since CRP is known to bend DNA sharply upon binding

(Schultz et al. 1991), we suggest that at many sites, it behaves more like a nucleoid-associated protein than a transcription factor.

Taken together, the available data argue that there is no intrinsic difference between nucleoid-associated proteins and transcription factors. As for all biological measurements, they are best considered in the light of evolution. Thus, it is possible that the need for bacteria to compact their genomic DNA came before the need to regulate transcription and that nucleoid proteins evolved as one of the first DNA binding proteins. As time passed, it is easy to believe that the genes encoding these proteins were duplicated, that sequence specificity evolved, and that cells that could regulate the transcription of certain loci were advantaged. Presumably, regulatory modules were grafted onto some of these factors, many of which then lost their function as nucleoid organizers. Prompted by data, such as that in Figs. 18.9 and 18.10, we suggest that *E. coli*, and probably many other bacteria, contain DNA binding proteins with a continuum of binding specificities and functions, and that the distinction between nucleoid-associated proteins and transcription factors is artificial.

18.8 Perspectives

E. coli is found in many places, and most of these, such as the guts of animals and aquatic environments, are subject to rapid and frequent fluctuations. As for most bacteria, survival depends on the selective expression of gene products to cope with the environment, and thus, it is no surprise that *E. coli* has evolved sophisticated systems to control transcription. This is most apparent in the high proportion of its gene products that are dedicated to regulating transcription initiation and in the complexity of even the simplest promoter. Thus, the *nir* operon promoter is regulated by four transcription factors: by FNR, by NarL (and its homologue, NarP) and by FruR and their activity is modulated by three nucleoid-associated proteins, IHF, FIS and H-NS. Although we can assume that different combinations of these factors are used in different conditions, most studies have been performed in ‘simple’ laboratory conditions, and the relative importance of the different factors in ‘real’ environments is still poorly understood. A quick glance at the *Ecocyc* database will convince anyone that the simplistic models for promoter regulation that appear in the textbooks are misleading. These models are mostly based on a small number of paradigm promoters (such as the *lac* promoter) and were established early in the history of this subject area. We now know that many, if not most, promoters are very complicated, with multiple factors interacting and other factors such as small ligands, the local chromosome landscape and DNA topology intervening. The challenge now is for us to put all the facts together, to produce integrated models, and, most important, to understand how systems are evolving.

Perhaps the most striking feature of transcriptional regulation in *E. coli* is its complexity, which, surely, has arisen from its evolution. Thus a simple DNA binding protein such as FIS, which has only 98 amino acids plays a myriad of roles

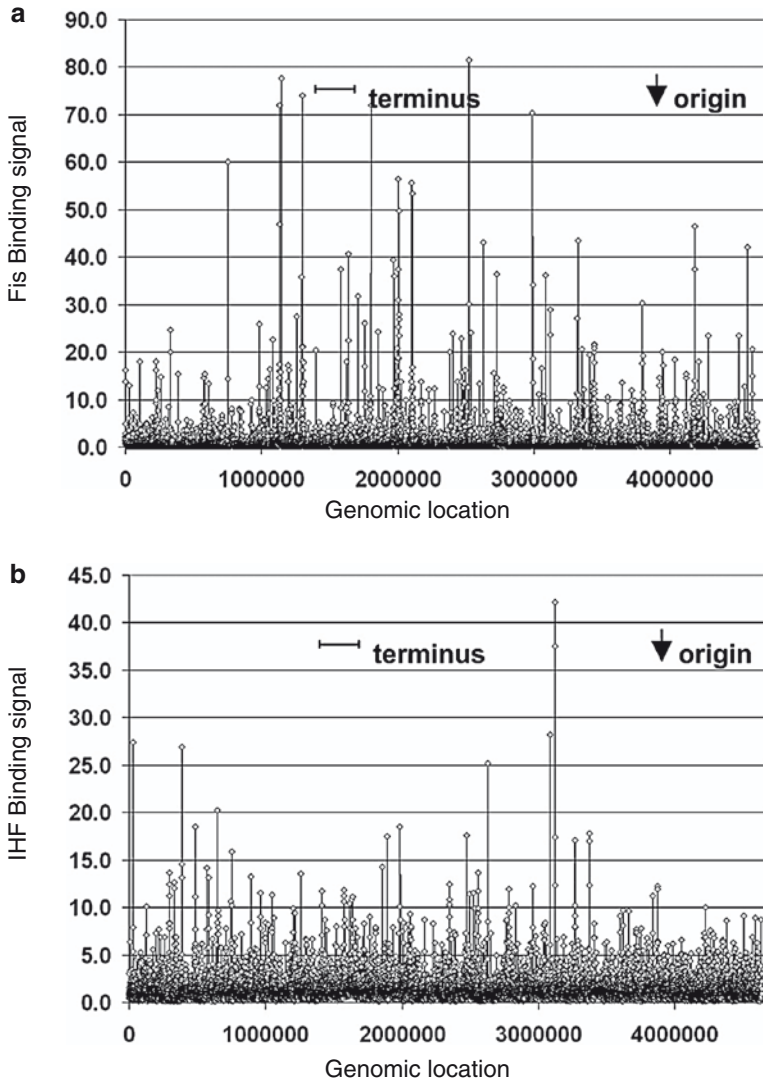


Fig. 18.9 Binding of FIS and IHF across the *E. coli* chromosome. The figure shows chromosome-wide DNA binding profiles for (a) FIS, and (b) IHF, generated from chromatin immunoprecipitation experiments in which immunoprecipitated DNA was analysed on high density microarrays (Grainger et al. 2006). The x-axes indicate sequence coordinates on the chromosome of *E. coli* K-12 strain MG1655 and the y-axes indicate the signal intensity, and hence the amount of FIS/IHF binding at that position

(Finkel and Johnson 1992). Presumably it originally evolved as a DNA binding protein with the ability to bend DNA. Its specificity then evolved to focus its binding at regulatory regions where it ‘learnt’ to participate in transcriptional regulation either as an activator or as a repressor. In our work, we have found 4 different ways

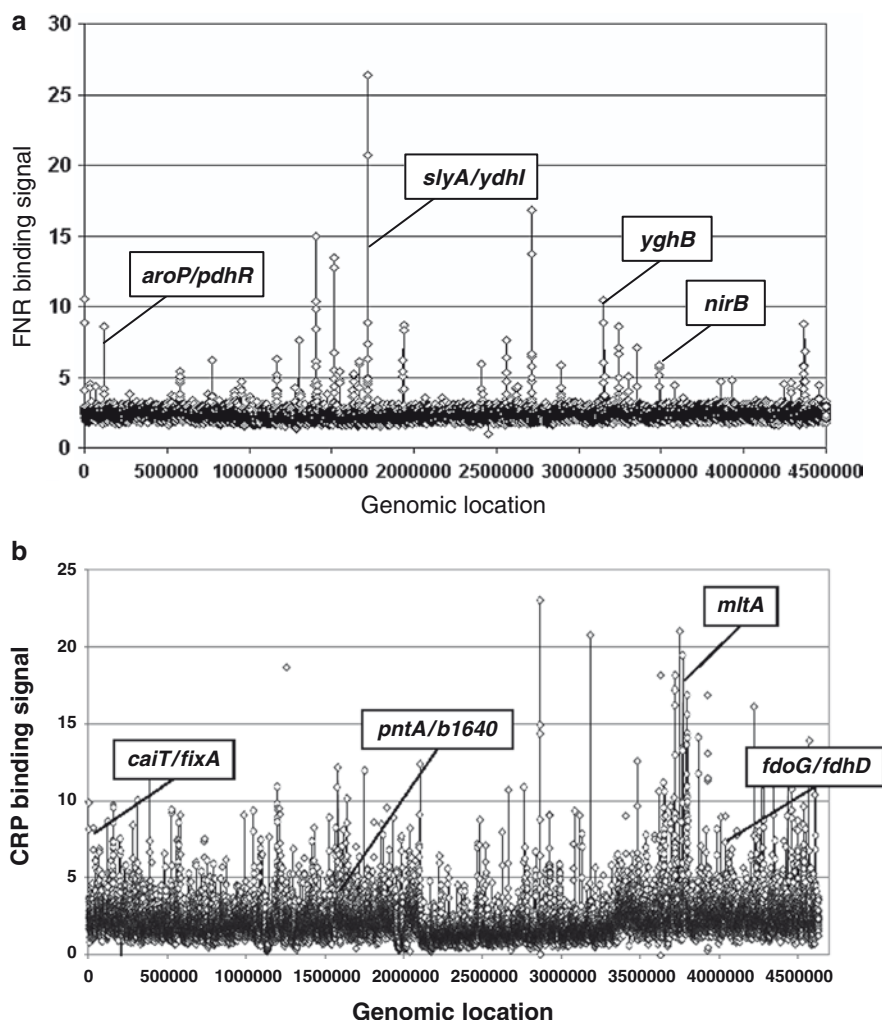


Fig. 18.10 Binding of FNR and CRP across the *E. coli* chromosome. The figure shows chromosome-wide DNA binding profiles for (a) FNR, and (b) CRP, generated from chromatin immunoprecipitation experiments in which immunoprecipitated DNA was analysed on high density microarrays (Grainger et al. 2005, 2007). The x-axes indicate sequence coordinates on the chromosome of *E. coli* K-12 strain MG1655 and the y-axes indicate the signal intensity, and hence the amount of transcription factor binding at that position. The locations of four targets in each profile are indicated

in which an *E. coli* promoter can be repressed by FIS. At the *nir* promoter, FIS helps IHF to suppress FNR-dependent activation. At the *nrf* promoter, FIS acts as a ‘simple’ repressor, whilst at the *ogt* promoter, it represses by displacing an essential

activator, NarL. Finally, at the *dps* promoter FIS acts as a sigma factor-dependent repressor of RNA polymerase holoenzyme containing σ^{70} . Presumably, these functions have been added gradually to the FIS repertoire and more are evolving. A recent chromatin immunoprecipitation study suggested that FIS may interact at nearly 900 targets in the *E. coli* genome (Cho et al. 2008). Our experience with CRP, FNR and RutR suggests that FIS binding at many of these targets will have little or no effect on the activity of neighbouring promoters but that targets provide a pool from which regulatory interactions can evolve. The arrival of high throughput rapid whole genome analytical strategies now provides a way to test these ideas.

Acknowledgements Work in the authors' laboratory has been supported by the Wellcome Trust and the UK BBSRC.

References

- Almiron M, Link A, Furlong D, Kolter R (1992) A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev* 6:1246–1254
- Altuvia S, Almiron M, Huisman G, Kolter R, Storz G (1994) The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol Microbiol* 13:265–272
- Ball C, Osuna R, Ferguson K, Johnson R (1992) Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J Bacteriol* 174:8043–8056
- Barnard A, Wolfe A, Busby S (2004) Regulation at complex bacterial promoters: how bacteria use different promoter organisations to produce different regulatory outcomes. *Curr Opin Microbiol* 7:102–108
- Beatty C, Browning D, Busby S, Wolfe A (2003) CRP-dependent activation of the *Escherichia coli* *acsP2* promoter by a synergistic Class III mechanism. *J Bacteriol* 185:5148–5157
- Bell A, Cole J, Busby S (1990) Molecular genetic analysis of an FNR-dependent anaerobically-inducible *Escherichia coli* promoter. *Mol Microbiol* 4:1753–1763
- Brown N, Stoyanov J, Kidd S, Hobman J (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* 27:145–163
- Browning D, Busby S (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2:57–65
- Browning D, Cole J, Busby S (2000) Suppression of FNR-dependent transcription activation at the *Escherichia coli* *nir* promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleoprotein assembly. *Mol Microbiol* 37:1258–1269
- Browning D, Lee D, Green J, Busby S (2002a) Secrets of bacterial transcription initiation taught by the *Escherichia coli* FNR protein. In: Hodgson D, Thomas C (eds) Signals, switches, regulons and cascades: control of bacterial gene expression. SGM Symposium, volume 61, pp 127–142
- Browning D, Beatty C, Wolfe A, Cole J, Busby S (2002b) Independent regulation of the divergent *nrfA* and *acsP1* *Escherichia coli* promoters by a nucleoprotein assembly at a shared regulatory region. *Mol Microbiol* 43:687–702
- Browning D, Cole J, Busby S (2004a) Transcription activation by remodeling of a nucleoprotein assembly: the role of NarL at the FNR-dependent *Escherichia coli* *nir* promoter. *Mol Microbiol* 53:203–215
- Browning D, Beatty C, Sanstad E, Gunn K, Busby S, Wolfe A (2004b) Modulation CRP-dependent transcription at the *Escherichia coli* *acsP2* promoter: repression by the nucleoid proteins FIS and IHF. *Mol Microbiol* 51:241–254
- Browning D, Grainger D, Beatty C, Wolfe A, Cole J, Busby S (2005) Integration of three signals at the *Escherichia coli* *nrf* promoter: a role for Fis protein in catabolite repression. *Mol Microbiol* 57:496–510

- Browning D, Lee D, Wolfe A, Cole J, Busby S (2006) The *Escherichia coli* K-12 NarL and NarP proteins insulate the *nrf* promoter from the effects of integration host factor. *J Bacteriol* 188:7449–7456
- Browning D, Cole J, Busby S (2008) Regulation by nucleoid-associated proteins at the *Escherichia coli nir* operon promoter. *J Bacteriol* 190:7258–7267
- Busby S, Ebright R (1994) Promoter structure, promoter recognition and transcription activation in prokaryotes. *Cell* 79:743–746
- Cho B-K, Knight E, Barrett C, Palsson B (2008) Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Res* 18:900–910
- Choy H, Adhya S (1996) Negative control. In: Neidhardt F (ed) *Escherichia coli* and *Salmonella*, vol 1. ASM Press, Washington, DC, pp 1287–1299
- Constantinidou C, Hobman J, Griffiths L, Patel M, Penn C, Cole J, Overton T (2006) A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrite, nitrate, NarXL, and NarQP as *Escherichia coli* K-12 adapts from aerobic to anaerobic growth. *J Biol Chem* 24:4802–4815
- Dame R (2005) The role of nucleoid-associated proteins in the organisation and compaction of bacterial chromatin. *Mol Microbiol* 56:858–870
- Finkel S, Johnson R (1992) The Fis protein: it's not just for DNA inversion anymore. *Mol Microbiol* 6:3257–3265
- Grainger D, Busby S (2008) Methods for studying global patterns of DNA binding by bacterial transcription factors and RNA polymerase. *Biochem Soc Trans* 36:754–757
- Grainger D, Hurd D, Harrison M, Holdstock J, Busby S (2005) Studies of the distribution of *Escherichia coli* cAMP receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci USA* 102:17693–17698
- Grainger D, Hurd D, Goldberg M, Busby S (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res* 34:4642–4652
- Grainger D, Aiba H, Hurd D, Browning D, Busby S (2007) Transcription factor distribution in *Escherichia coli*: studies with FNR protein. *Nucleic Acids Res* 35:269–278
- Grainger D, Goldberg M, Lee D, Busby S (2008) Selective repression by Fis and H-NS at the *E. coli dps* promoter. *Mol Microbiol* 68:1366–1377
- Harborne N, Griffiths L, Busby S, Cole J (1992) Transcriptional control, translation and function of the products of the five open reading frames of the *E. coli nir* operon. *Mol Microbiol* 6:2805–2813
- Helmann J, Chamberlin M (1988) Structure and function of bacterial sigma factors. *Ann Rev Biochem* 57:839–872
- Hollands K, Busby S, Lloyd G (2007) New targets for the cyclic AMP receptor protein in the *E. coli* K-12 genome. *FEMS Lett* 274:89–94
- Hussain H, Grove J, Griffiths L, Busby S, Cole J (1994) A seven gene operon essential for formate-dependent nitrite reduction to ammonia by enteric bacteria. *Mol Microbiol* 12:153–163
- Ishihama A (1997) Adaptation of gene expression in stationary phase bacteria. *Curr Opin Gen Dev* 7:582–588
- Jayaraman S, Gaston K, Cole J, Busby S (1988) The *nirB* promoter of *E. coli*: location of nucleotide sequences essential for regulation by oxygen, the FNR protein and nitrite. *Mol Microbiol* 2:527–530
- Karp P, Keseler I, Shearer A, Latendresse M, Krummenacker M, Paley S, Paulsen I, Collado-Vides J, Gama-Castro S, Peralta-Gil M, Santos-Zavaleta A, Penaloza-Spinola M, Bonavides-Martinez C, Ingraham J (2007) Multidimensional annotation of the *Escherichia coli* K-12 genome. *Nucleic Acids Res* 35:7577–7590
- Kolb A, Spassky A, Chapon C, Blazy B, Buc H (1983) On the different binding affinities of CRP at the *lac*, *gal* and *malT* promoter regions. *Nucleic Acids Res* 11:7833–7852
- Lin H, Bledsoe P, Stewart V (2007) Activation of *yeaR-yoaG* operon transcription by the nitrate-responsive regulator NarL is independent of oxygen-responsive regulator Fnr in *Escherichia coli* K-12. *J Bacteriol* 189:7539–7548
- Martinez-Antonio A, Collado-Vides J (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* 6:482–489

- McLeod S, Johnson R (2001) Control of transcription by nucleoid proteins. *Curr Opin Microbiol* 4:152–159
- Miroslavova N, Busby S (2006) Investigations of the modular structure of bacterial promoters. *Biochem Soc Symp* 73:1–10
- Murakami K, Darst S (2003) Bacterial RNA polymerases: the whole story. *Curr Opin Struct Biol* 13:31–39
- Ohniwa R, Morikawa K, Kim J, Ohta T, Ishihama A, Wada C, Takeyasu K (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. *EMBO J* 25:5591–5602
- Opel M, Aeling K, Holmes W, Johnson R, Benham C, Hatfield G (2004) Activation of transcription initiation from a stable RNA promoter by a Fis protein-mediated DNA structural transmission mechanism. *Mol Microbiol* 53:665–674
- Page L, Griffiths L, Cole J (1990) Differential physiological roles of two independent pathways for nitrite reduction to ammonia in enteric bacteria. *Arch Microbiol* 154:349–354
- Perez-Rueda E, Collado-Vides J (2000) The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. *Nucleic Acids Res* 28:1838–1847
- Potter P, Wilkinson M, Fitton J, Carr F, Brennan J, Cooper D, Margeson G (1987) Characterisation and nucleotide sequence of *ogt*, the O6-alkylguanine-DNA-alkyl transferase of *E. coli*. *Nucleic Acids Res* 15:9177–9193
- Rhodius V, Busby S (1998) Positive activation of gene expression. *Curr Opin Microbiol* 1:152–159
- Richet E, Vidal-Ingigliardi D, Raibaud O (1991) A new mechanism for coactivation of transcription initiation: repositioning of an activator triggered by the binding of a second activator. *Cell* 66:1185–1195
- Robison K, McGuire A, Church G (1998) A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete *Escherichia coli* K-12 genome. *J Mol Biol* 284:241–254
- Rojo F (2001) Mechanisms of transcriptional repression. *Curr Opin Microbiol* 4:145–151
- Roy S, Dinitriadi E, Kar S, Geanakopoulou M, Lewis M, Adhya S (2005) Gal repressor-operator-HU ternary complex: pathway of repressosome formation. *Biochemistry* 44:5373–5380
- Saier M, Ramseier T (1996) The catabolite repressor/activator protein of bacteria. *J Bacteriol* 178:3411–3417
- Samson L (1992) The suicidal DNA repair methyltransferases of bacteria. *Mol Microbiol* 6:825–831
- Schnetz K (2008) Fine-tuned growth control of *dps*, encoding a DNA protection protein, by FIS and H-NS. *Mol Microbiol* 68:1345–1347
- Schroder I, Darie S, Gunsalus R (1993) Activation of the *Escherichia coli* nitrate reductase (*narGHJI*) operon by NarL and FNR requires integration host factor. *J Biol Chem* 268:771–774
- Schultz S, Shields G, Steitz T (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* 253:1001–1007
- Sclavi B, Beatty C, Thach D, Fredericks C, Buckle M, Wolfe A (2007) The multiple roles of CRP at the complex *acs* promoter depend on activation region 2 and IHF. *Mol Microbiol* 65:425–440
- Scott S, Busby S, Beacham I (1995) Transcriptional co-activation at the *ansB* promoters: involvement of the activating regions of CRP and FNR when bound in tandem. *Mol Microbiol* 18:521–531
- Semsey S, Virnik K, Adhya S (2005) A gamut of loops: meandering DNA. *Trends Biochem Sci* 30:334–341
- Sheridan S, Opel M, Hatfield G (2001) Activation and repression of transcription initiation by a distant DNA structural transition. *Mol Microbiol* 40:684–690
- Shimada T, Fujita N, Maeda M, Ishima A (2005) Systematic search for the Cra-binding promoters using genomic SELEX system. *Genes Cells* 10:907–918
- Shimada T, Ishihama A, Busby S, Grainger D (2008) The *Escherichia coli* RutR transcription factor binds at targets within genes as well as intergenic regions. *Nucleic Acids Res* 36:3950–3955

- Talukder A, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181:6361–6370
- Taverna P, Sedgwick B (1996) Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli*. *J Bacteriol* 178:5105–5111
- Thanbichler M, Wang S, Shapiro L (2005) The bacterial nucleoid: a highly organized and dynamic structure. *J Cell Biochem* 96:506–521
- Tyson K, Bell A, Cole J, Busby S (1993) Identification of the nitrite and nitrate response elements at the anaerobically-inducible *Escherichia coli nirB* promoter. *Mol Microbiol* 7:151–157
- Tyson K, Cole J, Busby S (1994) Nitrite and nitrate regulation at the promoters of two *Escherichia coli* operons encoding nitrite reductase. *Mol Microbiol* 13:1045–1046
- Tyson K, Busby S, Cole J (1997) Catabolite regulation of two *E. coli* operons encoding nitrite reductases: role of Cra protein. *Arch Microbiol* 168:240–244
- Wang M, Gunsalus R (2000) The *nrfA* and *nirB* nitrite reductase operons in *Escherichia coli* are expressed differentially in response to nitrate than to nitrite. *J Bacteriol* 182:5812–5823
- Weiss B (2006) Evidence for mutagenesis by nitric oxide during nitrate metabolism in *Escherichia coli*. *J Bacteriol* 188:829–833
- Wigneshweraraj S, Bose D, Burrows P, Joly N, Schumacher J, Rappas M, Pape T, Zheng X, Stockley P, Severinov K, Buck M (2008) Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. *Mol Microbiol* 69:538–546
- Wing H, Williams S, Busby S (1995) Spacing requirements for transcription regulation by *E. coli* FNR protein. *J Bacteriol* 177:6704–6710
- Wolfe A (2005) The acetate switch. *Microbiol Mol Biol Rev* 69:12–50
- Wu H-C, Tyson K, Cole J, Busby S (1998) Regulation of the *E. coli nir operon* by two transcription factors: a new mechanism to account for co-dependence on two activators. *Mol Microbiol* 27:493–505

Index

A

Actin, 35–37, 44, 49, 51, 53, 58, 60–65, 87
Activator, 159, 282, 283, 285, 289, 332, 334,
338, 344, 356, 357, 363, 370, 371, 421,
423, 424, 426, 428, 430, 432, 433, 436,
440, 442
Alba, 212–215, 217
Anti-silencing, 253, 276–289, 297
Archaea, 3, 6, 39, 49, 50, 52, 159, 207–217,
355, 361, 362
AT-rich DNA, 158, 159, 259–261, 296,
383–386, 390, 391

B

Bacteria, 3, 6, 8, 14, 15, 17, 22, 24–27, 31–44,
63, 71, 72, 74, 75, 83, 84, 91, 97, 98,
101, 102, 104–108, 110–111, 120–122,
125, 128, 132, 136, 142, 152, 154, 162,
164, 167, 177, 179, 185–196, 207, 214,
217, 232, 234, 237, 247–250, 253–255,
263, 265, 270, 275, 283, 288, 293–296,
305, 306, 355–357, 362, 370, 371, 374,
422, 427, 439
Bacterial chromosome, 1, 3–8, 34, 52,
71, 72, 76, 77, 97–114, 119, 123,
130, 163, 182, 253, 326, 355, 358,
371, 423
bgl operon, 293, 297, 305, 306

C

Catabolite activator protein (CAP), 357
CC1, 210
Cell cycle, 17–20, 32–35, 39–41, 43, 60, 62,
65, 71, 72, 86, 87, 90, 91, 97, 98, 106,
112–113, 117
Chromatid, 53, 72, 73, 86
Chromatin remodelling, 78

Chromosome

arms, 25, 71, 72, 82, 84–85, 87, 91, 105
dynamics, 15, 31, 92, 125, 151, 182
segregation, 5, 6, 18, 31–44, 63, 72,
81–91, 97–99, 101, 102, 104–108,
110–111, 113, 117, 136
Cohesion, 24, 32, 33, 53, 73, 85–87, 91, 135
Condensins, 39, 72, 73, 122, 134, 135
Cren7, 209–211, 217
CRP, 249, 278, 328, 334, 337, 422, 423, 430,
432, 433, 435, 438, 441, 442
Crystallization, 100, 154, 180–182, 191, 194
Cytokinesis, 18–21, 49, 50, 107
Cytoplasm, 3, 7, 14, 17, 58, 74–77, 80, 90, 91,
108, 182, 190
Cytoskeleton, 36, 63

D

DNA

architectural proteins, 6, 247, 397
bending, 55, 123, 124, 157, 168, 262,
363, 367–392, 406, 424
bridging, 158, 168, 231, 249, 250, 260, 261,
263, 296, 304, 306, 332, 340, 358, 424
curvature, 119, 374
looping, 110, 119, 141, 165, 184, 284, 327,
358, 372, 397–401, 403–415, 422
protection, 26, 27, 166, 167, 178, 182, 192–194
supercoiling, 6, 7, 19, 25, 40, 118, 120,
130, 137, 152, 212, 247, 250, 274, 277,
288, 333, 341, 406, 415, 424
topology, 137, 151, 154, 157, 182, 250,
325, 327, 330, 332, 358, 374, 439
transactions, 164, 169, 247, 326, 327, 330,
336, 409
DnaA, 19, 137, 139, 169, 193, 332, 424
DNA adenine methyltransferase (Dam),
356, 357

Domainins, 141, 273

Domains, 4, 16, 24, 35, 38, 39, 55, 57, 59,
63, 77–80, 83, 88, 90, 92, 106, 118,
120–122, 125–132, 135, 140, 141,
152–162, 166, 184, 187, 207, 210–212,
214, 235, 249, 255–258, 260–262, 265,
268–270, 273–276, 286, 287, 292, 294,
296, 297, 302, 303, 305–307, 326, 327,
333, 335, 336, 340–345, 359–361, 369,
371, 376, 377, 399, 422

Dps, 26, 125, 151, 154, 166–169, 177–196,
231, 232, 234, 247, 249, 250, 271,
279–283, 340, 421, 424, 425, 432,
436–438, 442

E

Entropy, 90, 91, 98–102, 107, 108,
110–112, 263

Escherichia coli, 4, 14–17, 19–26, 32, 34,
35, 38, 39, 41, 44, 71–75, 77–82, 84,
85, 87, 88, 90–92, 97–99, 104–108,
110–113, 117, 119–121, 123–125,
127, 128, 130–133, 135–142, 150,
152, 154, 158–167, 178–186, 188–192,
194, 213, 231, 235, 236, 249, 250,
253–257, 263–268, 270–272, 274,
276, 277, 281–283, 285–287, 290,
292–296, 302, 304–306, 337–340,
342, 355, 356, 358–363, 368–376,
381, 397, 398, 401, 407, 421–442

Eukaryote, 6, 13, 24, 27, 39, 49, 50, 53, 65,
72, 73, 78, 87, 91, 119, 122, 207, 212,
214, 215, 223, 232, 234–236, 274, 275

Eukaryotic chromatin, 4, 141, 207, 211,
221–237

F

Factor for inversion stimulation (FIS), 76, 124,
141, 151–154, 158, 159, 162, 168, 222,
231, 234, 249, 250, 325–345, 368, 371,
401, 421, 424–442

Ferritin, 166, 167, 178, 179, 190, 192

Ferroxidase, 177–179, 184, 191

30nm Fibre, 226–231, 233, 234

Fluorescence, 6, 15–17, 20–23, 25, 62, 72, 74,
75, 78–80, 84, 86, 90, 92, 331, 364

FNR, 338, 423, 426–432, 438, 439, 441, 442

FtsK, 19, 21, 22, 25, 41, 43, 107, 108, 117,
122–123, 125, 135, 142

FtsZ, 18–21, 24, 34, 42–44, 107

Fur, 186–189

G

gal operon, 164, 165, 397, 398, 415

GalR, 165, 235, 397–401, 403–413,
415, 424

Genetic network, 168, 247, 328, 343

Gram negative, 120, 121, 154, 185, 186,
189, 270, 307, 355, 370, 380

Gram positive, 41, 43, 64, 186–189,
195, 254, 306, 355

Growth phase-dependent transcription,
326, 337

Gyrase, 38–40, 106, 107, 117–121, 129, 130,
135–137, 142, 163, 182, 294, 328, 332,
341, 342

H

Hha, 159, 253, 257, 266–271, 290, 291, 296,
302, 303, 308

High mobility group (HMG), 211, 222, 401

Histone-like nucleoid structuring (H-NS) protein,
26, 41, 76, 106, 124, 139–141,
151–159, 165, 168, 179, 185, 189,
191, 231, 232, 234–236, 249, 250,
253–308, 327–329, 340–342, 358–360,
368, 371, 372, 376, 377, 401, 424–428,
436, 437, 439

Histones, 5, 25, 123, 207, 209, 215–217,
222, 223, 229, 233, 254, 255,
274, 438

modification, 229, 233

octamer, 119, 215, 222, 223, 225, 229,
231–233

H₂O₂, 189–192

Horizontal gene transfer, 117, 142, 159,
253, 277

HU, 5, 25, 41, 76, 106, 123, 124, 151,
164–165, 168, 179, 217, 222, 231,
234–236, 249, 250, 278, 281–283,
294, 329, 332, 333, 363, 367–369,
371–379, 383, 387–392, 397–415,
424, 425

I

Integration host factor (IHF), 54–56, 76, 106,
119, 123, 158, 159, 163–165, 168, 185,
186, 231, 249, 250, 278, 279, 281–283,
285, 286, 292, 293, 328, 332, 335,
337–339, 367–392, 401, 407, 410,
423–430, 433, 435–441

Inversion, 124, 132, 152–154, 159, 165, 293,
298–301, 326, 332, 333, 371

L

LacI, 5, 17, 32, 279, 284, 285, 398–400, 404, 409, 410
 Lateral gene transfer, 217, 290–293
 Ler, 280, 285–287, 294, 296, 307
 Leucine, 159–163, 249, 269, 356–358, 360–363, 399
 Leucine-responsive regulatory protein (LRP), 152, 158–163, 168, 231, 249, 250, 278, 290, 293, 301, 339, 355–364, 424
 Linker histone, 222, 228, 229, 233
 Lsr2, 305–306

M

Macromolecular crowding, 7, 26, 75, 415, 424
 MC1, 211, 216, 217
 Methylation, 209, 211, 233, 274, 356, 432
 Motor protein, 34, 36, 108
mreB, 35, 87
 MukBEF, 24, 25, 85, 106, 119, 122, 134–136, 141, 142, 294
 MvaT, 294, 304–305
 MvaU, 294, 304–305
Mycobacterium, 121, 178, 179, 183, 192, 196, 305

N

Nucleoid-associated protein (NAP), 7, 8, 25–27, 106, 119, 123–125, 132, 141, 142, 151–169, 179, 180, 182–186, 190, 222, 231, 232, 248–250, 254, 274, 279, 281, 327, 328, 330, 340, 343, 344, 355–364, 368, 371–373, 415, 421–442
 Nucleoids, 4–8, 13–27, 38, 39, 41, 42, 49–65, 71–92, 104–108, 119, 120, 125–128, 132–138, 140–142, 151–169, 179, 182, 185, 190, 191, 209, 215, 217, 221, 222, 232, 234, 236, 247–251, 254, 274, 278–282, 285, 288, 292, 293, 295, 296, 325–345, 355–364, 368, 371–373, 397, 400, 402, 414–415, 421–442
 segregation, 49–65, 74
 structure, 5, 71–92
 Nucleosome, 5, 215, 223–231, 233, 234, 415
 Nucleus, 3, 6, 13, 73, 222, 232

O

OriC, 14, 16, 17, 81–83, 86, 87, 97, 104, 105, 115, 131–137, 139, 140, 142, 169, 193, 331–332, 336

Origin, 6, 16, 17, 32–37, 52, 59, 72, 83–86, 90, 91, 97, 135, 139, 193, 326, 332, 344, 368, 401
 Oxidative stress, 26, 166, 177–179, 182, 185, 186, 188, 189, 192–195, 250, 282, 283, 302, 436, 437
 OxyR, 185, 186, 189, 279, 283, 308, 339, 436, 437

P

ParA, 36–37, 44, 51–53, 57–61, 64, 65
 ParB, 37, 42, 52, 53
 Partition, 50, 51, 53, 55, 60, 62–65
 Phase separation, 7, 74–77, 91, 92, 99, 100, 108
 Plasmid, 17, 36, 37, 44, 49–65, 80, 120, 121, 123, 124, 130, 137, 142, 149, 189, 264, 266, 276, 284, 289, 290, 294, 297, 300, 302–304, 307, 332, 370, 401, 430, 431
 Plasmid segregation, 49–65, 121
 Polymer physics, 92, 97–114
 ppGpp, 24, 163, 278, 328, 359
 Promoter, 56, 128, 130, 131, 153, 154, 158, 169, 185–189, 225, 231, 235, 236, 248, 249, 259–261, 263, 264, 268, 271–286, 288–290, 292, 293, 297–301, 303, 305, 306, 308, 334–338, 345, 356–359, 370, 397, 400, 402, 404, 422–424, 426–439, 441, 442
proU, 264, 267, 271, 273–280, 287, 297–302, 307, 308
Pseudomonas putida, 52, 192, 363, 369, 370, 374–377, 379, 381, 383, 385, 386

R

Recombination, 5, 22, 41, 42, 50, 118, 123, 126, 127, 129, 130, 135, 136, 139, 142, 152, 159, 163, 169, 224, 225, 232, 290, 293, 326, 330, 332, 333, 336, 368, 373, 402
 Replication, 16–19, 21, 26, 32–35, 38–41, 50, 64, 72, 73, 76, 77, 81, 83–87, 90–92, 97, 98, 104, 105, 112, 118, 120, 121, 126, 132–139, 142, 149, 152, 163, 169, 193, 196, 215, 232, 262, 288, 292, 295, 326, 330–332, 336, 340, 344, 368, 401, 402
 Replichores, 73, 77, 82, 84–85, 90, 91, 135, 342
 Repressor, 16, 32, 51, 53, 57, 79, 84, 87, 128, 129, 156, 159, 186, 225, 235, 236, 250, 306, 331, 337, 338, 340, 356, 398, 399, 422, 423, 426, 428, 430, 432, 436, 440–442

Repressosome, 165, 397, 399, 411, 412
 RNA polymerase, 22, 36, 119, 121, 126,
 128–132, 136, 153, 154, 158, 159, 185,
 186, 231, 236, 254, 280–281, 328, 338,
 344, 356, 357, 370, 397, 400, 414, 415,
 422–424, 428, 432, 436, 437, 442
 rRNA, 23, 24, 124, 126, 131, 132, 136, 150,
 153, 158, 282, 333, 358

S

Sac10a, 212
 Sac10b, 212
Salmonella, 123–128, 131, 132, 136–142, 156,
 188, 190, 195, 254, 256, 257, 263–267,
 270–272, 275–277, 282, 284, 289, 292,
 295, 296, 302–304, 339, 340
 SASP, 183, 191
 Segregation, 5, 6, 18, 26, 31–44, 49–65,
 71–92, 97–114, 120, 121, 126,
 134–137, 139, 142, 382
 σ^{38} , 272, 422, 436, 437
 σ^{70} , 23, 185, 186, 188, 189, 272, 330, 336,
 338, 370, 422, 436, 437, 442
 σ factor, 422, 436
 Sfh, 294, 302–304
 Silencing, 78, 154, 159, 235, 237, 254,
 265–267, 270–289, 292, 294, 295, 297,
 301, 302, 305, 307, 308, 340, 436
 Sir2, 214
 Site-specific recombination, 50, 118, 123, 130,
 293, 326, 330, 336, 368
spoIIIE, 19, 22, 25, 41, 43, 107, 108
 Stable RNA, 80, 153, 326, 333–336, 340, 344
 StpA, 119, 124, 159, 165, 265, 268, 270,
 294–297, 301–303, 307, 329,
 339, 424
 Structural maintenance of chromosomes
 (SMC), 24–26, 34, 39–42, 44, 106,
 119, 122, 222, 276, 294
 Sul7, 208–212, 216
 Supercoil diffusion, 125–128, 141
 Supercoiling, 6, 7, 19, 25, 38–40, 107, 111,
 118–121, 123–126, 130, 136–141, 152,
 169, 212, 234, 236, 247, 250, 261, 262,
 273–274, 277, 281, 288, 293, 298, 300,
 301, 332–334, 337, 340–345, 406, 411,
 413, 415, 424
 Superhelical density, 38, 234, 330, 332, 334,
 340, 344, 413

Superhelicity, 76, 78, 81, 182, 189, 222, 232,
 234, 277, 328–330, 332–334, 336, 337,
 339–344, 404

T

Terminus, 6, 16, 21, 22, 32, 34, 41, 84, 97,
 132–136, 157, 167, 214, 256–257, 340,
 344, 360, 399
TFI, 372–379, 383, 384, 387–392
 Topoisomerase, 18, 19, 25, 37, 87, 108,
 120–123, 142, 182, 340, 341
 Topoisomerase I (Topo I), 37–40, 119–121,
 123, 124, 182
 Topo IV, 38–40, 87, 108, 119–121, 123, 135,
 137, 142
 Transcription, 7, 22–24, 36, 38, 53, 76, 77,
 80–81, 84, 85, 87, 88, 91, 92, 98, 118,
 120, 124, 128–132, 138, 140, 141, 152,
 153, 158, 159, 162–165, 167–169,
 185–189, 212, 214, 215, 222, 231–236,
 247–250, 254, 256, 258, 260–262, 264,
 271–274, 276, 278–280, 282–287, 293,
 298–300, 302–304, 308, 326–330,
 333–345, 356–358, 360, 362, 363, 368,
 370, 372, 373, 397–400, 402, 404, 406,
 410, 413–415, 421–441
 initiation, 189, 330, 333, 334, 338, 340,
 344, 363, 415, 422, 424, 426, 427, 429,
 438, 439
 regulation, 7, 158, 162–164, 169, 250, 397,
 422–423, 427, 429, 433, 435
 tRNA, 142, 153, 254
 Tubulin, 34, 51, 63, 64

V

virF, 263, 266, 271, 276, 282, 289
 Virulence, 50, 154, 158, 189, 193–196, 254,
 263–266, 276, 277, 282, 288, 289, 292,
 293, 304, 339, 340, 373
 Volume-exclusion, 84

X

Xenogeneic silencing, 265, 274, 288, 295

Y

YdgT, 266–268, 270, 296